

ISSN 1825-0211



**UN'ESTENSIONE STOCASTICA  
DEL MODELLO "FISHER-LANGE"**

**Massimo DE FELICE — Franco MORICONI**

**Quaderno n. 86 — Marzo 2011**

**QUADERNI DEL DIPARTIMENTO  
DI ECONOMIA, FINANZA  
E STATISTICA**

---

# Un'estensione stocastica del modello “Fisher-Lange”

Massimo De Felice, “Sapienza” - Università di Roma

Franco Moriconi, Università di Perugia

## Sunto

Tra i metodi di stima puntuale della riserva sinistri nell'assicurazione danni è largamente utilizzato, in particolare dalle compagnie italiane, un approccio noto come “metodo di Fisher-Lange” (FL). L'FL è un metodo “a costi medi”, nel senso che la valutazione degli impegni di rimborso sinistri è scomposta in una stima del costo medio del singolo sinistro e in una stima del numero di sinistri da rimborsare. Una seconda caratteristica dell'FL è che la stima del numero di sinistri da pagare è basata su una *case outstanding development technique*, in cui lo sviluppo del numero di sinistri da pagare ha come variabile esplicativa dati da riserva di inventario (*case reserves*). L'FL è un metodo essenzialmente deterministico; si propone qui un “modello FL stocastico” (SFL), cioè una versione probabilistica del metodo FL. L'approccio stocastico consente di costruire l'intera distribuzione di probabilità dei costi di rimborso futuri, producendo quindi una adeguata misurazione del *reserve risk* e una quantificazione del corrispondente requisito patrimoniale, come richiesto ai modelli interni nell'ambito della Direttiva Quadro *Solvency II*. Una caratteristica saliente del modello SFL è quella di fornire una estensione stocastica coerente dell'approccio deterministico, nel senso che gli stimatori per i valori attesi prodotti dall'SFL coincidono con gli stimatori empirici adottati nella procedura di stima puntuale, e sono dotati delle qualità statistiche necessarie per garantire l'affidabilità previsiva (non distorsione, varianza minima).

Il modello proposto ha struttura trivariata, nel senso che lo sviluppo dei costi è determinato da tre processi stocastici: il processo del costo medio, il processo del numero di sinistri “con seguito” e il processo del numero di sinistri pagati. L'informazione *case outstanding* agisce sui processi del numero, che hanno come variabile esplicativa il numero dei sinistri messi a riserva in ogni anno di sviluppo. La struttura dei processi ipotizzati consente di costruire per simulazione la distribuzione di probabilità di tutti gli impegni futuri, sia secondo l'approccio *Liability-at-Maturity*, sia secondo l'approccio *Year-End Expectation*. L'incertezza di stima dei parametri è inclusa nelle distribuzioni effettuando le simulazioni con un approccio *conditional parametric bootstrap*. Vengono anche ricavate espressioni in forma chiusa per il *mean square error of prediction* (MSEP) dei costi di rimborso. Una approssimazione lineare di queste formule consente di riottenere come caso particolare, reinterpretando opportunamente le quantità coinvolte, le formule di MSEP ricavate in un modello di *loss reserving* recentemente proposto da Dahms.

*Parole chiave:* stochastic loss reserving, average cost, case outstanding development technique, mean square error of prediction, conditional parametric bootstrap.

*JEL classification:* C13, C15, C32

## 1 Introduzione

Tra i metodi di stima della riserva sinistri nell'assicurazione danni hanno particolare diffusione gli approcci “a costi medi”, in cui la valutazione degli impegni di rimborso per i sinistri delle

generazioni chiuse è scomposta in una stima del costo medio del singolo sinistro (*average cost, average payment per claim finalized*) e in una stima del numero di sinistri da rimborsare. Per ambedue le componenti sono rilevanti le ipotesi sulla forma dello “sviluppo”, cioè sulla legge di variazione in funzione dell’antidurata. Nelle applicazioni pratiche l’ipotesi più diffusa per gli importi pagati è che il costo medio, al netto degli effetti inflattivi, è costante per tutte le generazioni e cambia solo con l’anno di sviluppo. Per la componente dei numeri si incontra una maggiore varietà di ipotesi modellistiche. Una di queste consiste nell’adottare uno schema di tipo chain ladder, in cui la variabile esplicativa per lo sviluppo dei numeri futuri è il valore cumulato dei numeri osservati (modelli *age-to-age*). Una classe alternativa di modelli per i numeri è quella basata sulla *case outstanding development technique*, in cui lo sviluppo dei numeri futuri ha come variabile esplicativa dati da riserva di inventario (*case reserves*). In questo approccio un sinistro viene seguito fin dall’attribuzione di una riserva iniziale nell’istante di denuncia, ai successivi aggiustamenti dovuti a eventuali pagamenti parziali, fino alla definitiva liquidazione.

Si considera qui un approccio appartenente a questa seconda classe, largamente utilizzato nella pratica – in particolare dalle compagnie italiane – noto come “metodo di Fisher-Lange” (FL); la denominazione fa riferimento a una metodologia presentata nel lavoro [5], cui l’approccio FL si ispira, pur discostandosene però nell’effettiva realizzazione. Il metodo FL tradizionale è essenzialmente deterministico; rientra cioè tra i metodi di stima “puntuale”, in cui vengono utilizzati stimatori empirici per la proiezione dei valori (numeri e/o importi) futuri, senza produrre nessuna valutazione di variabilità e, più in generale, nessuna quantificazione dell’incertezza caratteristica di questi valori.

Si propone qui un “modello FL stocastico” (SFL), cioè una versione probabilistica del metodo FL deterministico. Una caratteristica saliente del modello è quella di fornire una *estensione stocastica coerente* dell’approccio deterministico, nel senso che gli stimatori per i valori attesi prodotti dall’SFL coincidono con gli stimatori empirici adottati nella procedura classica di stima puntuale, cioè nell’FL. Per ottenere questa proprietà il modello stocastico viene “costruito intorno” allo schema deterministico: gli stimatori dell’FL vengono adottati come punto di partenza, e la struttura probabilistica delle due componenti dell’SFL, numeri e costi medi, è scelta in modo che gli stimatori FL risultino avere le qualità statistiche necessarie per garantire l’affidabilità previsiva. Conseguenza rilevante di questa proprietà di coerenza è che una “validazione” dell’FL implica la “validazione” dell’SFL, e viceversa.

Vale la pena di segnalare che è stato recentemente proposto da Dahms [3] un modello per molti aspetti affine all’SFL qui presentato. Il modello di Dahms è anch’esso basato su una tecnica “case outstanding” – etichettata come “Complementary Loss Ratio Method” (CLRM) –, ma non usa un approccio a costi medi, le informazioni su pagati e riservati essendo riferite direttamente agli importi. Dal punto di vista della struttura matematica, il modello di Dahms e la componente per i numeri del’SFL presentano forti analogie formali. In particolare l’SFL per i numeri può essere inteso come una estensione “time series” del modello di Dahms, che è di tipo “distribution free”. L’analogia cessa però di valere se si considera l’SFL completo, dato che l’approccio a costi medi, espressivo se applicato agli importi pagati, perde di significatività se applicato agli importi riservati. I due approcci quindi, intesi come modelli previsivi per gli importi, risultano ispirati a una logica differente.

## 2 Definizioni preliminari

### 2.1 L’organizzazione per AY/DY

Si faccia riferimento a un ramo fissato, con dati organizzati per *accident year* (AY) e *development year* (DY).

Alla data corrente  $t$  si sia osservato un numero  $I$  di AY  $i = 1, \dots, I$ , per cui si può porre  $t = I$ .

Si può avere un numero  $J$  di DY  $j = 1, \dots, J$ . Per non appesantire l'esposizione si farà riferimento al caso:

$$I = J;$$

i dati relativi agli anni  $(i, j)$  trascorsi sono quindi in forma di “triangolo”. L'estensione al caso  $I \geq J$  (“trapezoide”) è comunque immediata.

Per ogni  $i$ , converrà definire l'“indice diagonale”  $d_i := I - i + 1$ .

## 2.2 Importi a costi correnti

Si assumerà che tutti gli importi monetari siano espressi a costi correnti, siano cioè rivalutati fino a  $t = I$  per l'inflazione di ramo.

## 2.3 Le quantità rilevanti

Per ogni  $(i, j)$ , cioè per  $1 \leq i, j \leq I$ , sia:

$C_{i,j}$  : l'ammontare degli importi pagati nel corso dell'anno  $(i, j)$ ,

$N_{i,j}$  : il numero di sinistri pagati chiusi nel corso dell'anno  $(i, j)$ ,

$R_{i,j}$  : il numero dei sinistri messi a riserva alla fine dell'anno  $(i, j)$  (al netto degli IBNyR<sup>1</sup>,

$D_{i,j}^{(1)}$  : il numero dei sinistri denunciati in  $(i, j)$  (senza gli IBNyR),

$D_{i,j}^{(2)}$  : il numero dei sinistri riaperti in  $(i, j)$ ,

$D_{i,j}^{(3)}$  : il numero dei sinistri chiusi senza seguito in  $(i, j)$ .

Si può subito osservare che, con queste definizioni, per ogni  $(i, j)$  vale la relazione di accumulo:

$$R_{i,j} = \sum_{k=1}^j (D_{i,k}^{(1)} + D_{i,k}^{(2)} - D_{i,k}^{(3)} - N_{i,k}).$$

*Osservazione.* Il fatto che gli importi  $C_{i,j}$  siano espressi a costi correnti sottintende che i dati sui costi storici  $C_{i,j}^{(h)}$  siano stati preliminarmente rivalutati sulla base di un modello per l'inflazione che caratterizza il ramo considerato. L'inflazione di ramo si può intendere come la somma dell'inflazione economica e dell'inflazione aggiuntiva specifica del ramo (*superimposed inflation*). Tipicamente si assume che l'inflazione di ramo agisca come un “effetto diagonale”, cioè che dipenda solo dalla distanza di tempo intercorrente tra l'anno di calendario corrente  $I$  e l'anno di calendario  $y(i, j) := i + j - 1$  cui si riferisce il pagamento  $C_{i,j}^{(h)}$ . Sotto questa ipotesi, quindi, la rivalutazione dei costi storici può essere espressa nella forma:

$$C_{i,j} = C_{i,j}^{(h)} e^{r_{y(i,j)} [I - y(i,j)]}, \quad i + j \leq I + 1,$$

dove  $r_{y(i,j)}$ , il tasso di inflazione realizzata tra gli anni di calendario  $y(i, j)$  e  $y(i, j) + 1$ , può anche essere imposto uguale a un valore costante  $r$  (inflazione “media” sul periodo da  $t = 1$  a  $t = I$ ). Si può osservare però che l'ipotesi di effetto diagonale è del tutto spontanea solo relativamente all'inflazione economica, mentre per la componente di inflazione specifica si potrebbe ipotizzare, per esempio, un trend diverso per ciascun anno di sviluppo. Si potrebbe cioè assumere il modello:

$$C_{i,j} = C_{i,j}^{(h)} e^{(r^{(e)} + r_j^{(l)}) [I - y(i,j)]}, \quad i + j \leq I + 1,$$

---

<sup>1</sup>Con la dicitura IBNyR (*Incurred But Not yet Reported*) si intendono qui i sinistri tardivi non ancora realizzati nell'anno  $(i, j)$ .

dove  $r^{(e)}$  è il tasso medio di inflazione economica e  $r_j^{(l)}$  è il tasso medio di inflazione specifica che caratterizza i costi di rimborso nel DY  $j$ . ■

## 2.4 La struttura dei dati

Si assumerà che l'insieme dei dati disponibili alla data  $t = I$ , cioè l'insieme delle terne  $\{C_{i,j}, N_{i,j}, R_{i,j}\}$  osservate<sup>2</sup> per  $i + j \leq I + 1$ , possa essere incompleto. Per garantire la validità di tutte le definizioni adottate nel seguito, converrà limitare il tipo di incompletezza ammissibile a dati con struttura convessa<sup>3</sup>. Sia:

$$Y_{i,j} := \{C_{i,j}, N_{i,j}, R_{i,j}\};$$

e, per  $1 \leq j \leq I$ , si definisca:

$$\text{may}(j) : \text{minimo } AY \text{ } i \text{ nel quale si ha un'osservazione di } Y_{i,j} \text{ nel DY } j,$$

e:

$$\mathcal{I}_j := \{i \mid \text{may}(j) \leq i \leq d_j\}.$$

Si assumeranno le proprietà, per  $1 \leq j \leq I$ :

$$\begin{aligned} Y_{i,j} \text{ è osservato per ogni } i \in \mathcal{I}_j; \\ 1 \leq \text{may}(j) \leq d_j; \\ \text{may}(j+1) \leq \text{may}(j), \quad 1 \leq j \leq I-1. \end{aligned} \tag{1}$$

Allora, definendo, per  $1 \leq i \leq I$ :

$$\text{mdy}(i) = \min\{j \mid \text{may}(j) \leq i\},$$

e:

$$\mathcal{J}_i := \{j \mid \text{mdy}(i) \leq j \leq d_i\},$$

varranno anche le proprietà:

$$\begin{aligned} Y_{i,j} \text{ è osservato per ogni } j \in \mathcal{J}_i; \\ 1 \leq \text{mdy}(i) \leq d_i; \\ \text{mdy}(i+1) \leq \text{mdy}(i), \quad 1 \leq i \leq I-1. \end{aligned}$$

L'insieme delle terne  $\{C_{i,j}, N_{i,j}, R_{i,j}\}$  osservate in  $t = I$  è dato da:

$$\mathcal{D}_I := \bigcup_{j=1}^I \{Y_{i,j} \mid i \in \mathcal{I}_j\};$$

date le proprietà (1),  $\mathcal{D}_I$  è un insieme convesso. Si assumerà che tutti gli  $Y_{i,j} \in \mathcal{D}_I$  abbiano elementi positivi, cioè che tutti i valori osservati di  $C_{i,j}, N_{i,j}, R_{i,j}$  siano maggiori di zero.

Si definiscono anche gli insiemi, per  $1 \leq j \leq I$ :

$$\mathcal{B}_j^R := \bigcup_{k=1}^j \{R_{i,k} \mid i \in \mathcal{I}_j\}, \quad \mathcal{B}_j^N := \bigcup_{k=1}^j \{N_{i,k} \mid i \in \mathcal{I}_j\};$$

<sup>2</sup>Con “osservate” si può intendere, in senso più esteso, “ritenute osservazioni utilizzabili a fini previsivi”.

<sup>3</sup>La convessità dell'insieme delle osservazioni viene richiesta per garantire la significatività delle ipotesi del modello anche qualora queste siano intese “in senso forte”, cioè siano assunte valide non solo per le variabili aleatorie in quanto tali, ma anche per i loro valori effettivamente osservati. L'ipotesi potrebbe essere rimossa con qualche accortezza (assumendo valide le ipotesi di modello indipendentemente dall'osservabilità delle variabili aleatorie coinvolte).

$\mathcal{B}_j^R$  e  $\mathcal{B}_j^N$  contengono le osservazioni dei numeri riservati e, rispettivamente, di numeri chiusi fino al DY  $j$ . Le osservazioni complessive sui numeri fino al DY  $j$  sono espresse da  $\mathcal{B}_j := \mathcal{B}_j^R \cup \mathcal{B}_j^N$ .

Il “triangolo” dei valori futuri sarà rappresentato dall’insieme:

$$\mathcal{F}_I := \{Y_{i,j} \mid i + j > I + 1\}.$$

## 2.5 Obiettivi del modello

In  $t = I$  il cuore del problema di loss reserving stocastico è la caratterizzazione dei pagati futuri  $C_{i,j} \in \mathcal{F}_I$ ; dato che le  $C_{i,j}$  future sono variabili aleatorie, si tratta di ricavarne la distribuzione di probabilità, o almeno di ottenerne una stima della media e della varianza. La distribuzione di probabilità, e quindi la varianza, dovrà contenere sia l’incertezza “intrinseca” dei pagati futuri, sia l’incertezza dovuta alla stima dei parametri del modello di valutazione. In termini di varianza, si tratterà di ricavare il *mean square error of prediction* (MSEP) delle  $C_{i,j}$ , comprensivo sia del *process error* che dell’*estimation error*.

Questo risultato cruciale andrà poi arricchito passando a caratterizzare, sempre in termini di distribuzione di probabilità, o almeno in termini di media e varianza, le seguenti variabili aleatorie “derivate”.

- I pagamenti  $C_{i,j}^*$  inflazionati, cioè gli importi  $C_{i,j}$  rivalutati con un opportuno modello di inflazione.
- I valori scontati  $M_{i,j}$ , cioè gli importi  $C_{i,j}^*$  attualizzati coi fattori di sconto risk-free in vigore sul mercato in  $t = I$  per la durata  $\tau = j - d_i$ .
- I *costi ultimi* di generazione  $U_i := \sum_{j=1}^I C_{i,j}^*$  e i costi ultimi totali  $U := \sum_{i=2}^I U_i$ . La parte dei costi ultimi non ancora pagata,  $L_i := \sum_{j=d_i+1}^I C_{i,j}^*$  e  $L := \sum_{i=2}^I L_i$ , caratterizza gli *impegni in-essere* (*Outstanding Loss Liabilities*, OLL) secondo l’approccio “*Liability-at-Maturity* (LM) non scontato”. In particolare il valore atteso  $\bar{L}_i := \mathbf{E}(L_i | \mathcal{D}_I)$ , e il corrispondente totale su tutte le generazioni aperte  $\bar{L} := \mathbf{E}(L | \mathcal{D}_I)$ , forniscono le stime di riserva LM non scontata. Le corrispondenti stime del capitale di sicurezza (*Risk Capital*, *Solvency Capital Requirement*), sempre secondo l’approccio LM, si ricavano definendo le opportune misure di variabilità.
- Le OLL scontate di generazione  $M_i := \sum_{j=d_i+1}^I M_{i,j}$  e le OLL totali scontate  $M := \sum_{i=2}^I M_i$ . Si ottiene in questo caso la caratterizzazione probabilistica degli impegni in-essere nell’approccio “LM scontato”. In particolare, il valore atteso  $V_i := \mathbf{E}(M_i | \mathcal{D}_I)$  per AY e il valore atteso totale  $V := \mathbf{E}(M | \mathcal{D}_I)$  di questi impegni produce le *Best Estimate*, che sono le stime di riserva coerenti con Solvency II (secondo l’approccio LM).
- La stima di riserva  $\hat{V}'$  prodotta “a fine anno”, cioè in  $t = I + 1$ , dalla quale si ricavano gli “obblighi di fine anno”  $Z := \hat{V}' + \sum_{i=2}^I C_{i,d_i+1}^*$ . Dalla distribuzione del valore attuale di  $Z$  (valore scontato col tasso risk-free a un anno) si ricavano stime del capitale di sicurezza coerente con Solvency II secondo l’approccio cosiddetto YEE (*Year-End Expectation*) [4], o CDR (*Claims Development Result*) [11].

## 3 Il modello per i costi medi

### 3.1 L’approccio a costi medi agli importi pagati

L’obiettivo principale è quindi quello di costruire un modello stocastico per le v.a.  $C_{i,j}$ , cioè un modello “previsivo” che fornisca la distribuzione di probabilità congiunta degli importi pagati futuri, o almeno una stima della media e della varianza di queste distribuzioni.

A tale scopo si adotterà un modello “frequency-severity”, che esprime il costo di rimborso nell’anno  $(i, j)$  come il prodotto:

$$C_{i,j} = K_{i,j} N_{i,j},$$

dove  $K_{i,j}$  è una v.a. che esprime il costo medio di un sinistro nell’anno  $(i, j)$ . I due fattori che formano il costo vanno descritti separatamente. L’approccio richiede cioè di specificare un modello previsivo per lo sviluppo della “severity”  $K_{i,j}$  e un modello previsivo per lo sviluppo della “frequency”  $N_{i,j}$ , che si assumerà essere positiva.

In questo paragrafo si caratterizzerà la componente severity, cioè il modello per i costi medi. Nel par. 4 verranno considerati i dettagli della componente frequency, quindi del modello per i numeri, e verranno riunite le due componenti nel modello globale per gli importi.

### 3.2 I costi medi unitari

Per ogni anno  $(i, j)$  si definiscano i *costi medi unitari* (“average payments per claim”):

$$K_{i,j} := \frac{C_{i,j}}{N_{i,j}}. \quad (2)$$

*Osservazione.* In realtà l’interpretazione della (2) come costo medio contiene una incoerenza. Infatti, mentre il denominatore  $N_{i,j}$  conta il numero di pagati chiusi, il numeratore  $C_{i,j}$  totalizza gli importi pagati a qualsiasi titolo. Oltre ai pagamenti per sinistri effettivamente chiusi in  $(i, j)$ , include quindi anche i pagamenti parziali in  $(i, j)$  per sinistri con seguito, mentre esclude gli eventuali pagamenti parziali per i chiusi in  $(i, j)$  effettuati negli anni precedenti. L’interpretazione della (2) è quindi completamente accettabile solo se si suppone che i due effetti di inclusione/esclusione si compensino.

L’incongruenza, dovuta alla difficile reperibilità dei dati corretti, è comunque riconosciuta e accettata in larga parte della letteratura. Per una discussione di questa problematica si veda [9], pp.87, 94-99. ■

### 3.3 Il modello deterministico per i costi medi

Lo schema FL deterministico per i costi medi<sup>4</sup> consiste nel supporre che in ciascun anno di sviluppo il costo medio è un parametro  $\kappa_j$  caratteristico del DY e non dipendente dall’AY. Per  $1 \leq i, j \leq I$  vale quindi la relazione di proporzionalità:

$$C_{i,j} = \kappa_j N_{i,j}.$$

Nella pratica attuariale corrente, alla data  $t = I$ , e quindi dato  $\mathcal{D}_I$ , si adotta di solito per  $\kappa_j$  lo stimatore:

$$\hat{\kappa}_j^{\text{FL}} := \frac{\sum_{i \in \mathcal{I}_j} C_{i,j}}{\sum_{i \in \mathcal{I}_j} N_{i,j}} = \sum_{i \in \mathcal{I}_j} \frac{N_{i,j}}{\sum_{h \in \mathcal{I}_j} N_{h,j}} K_{i,j}, \quad j = 1, \dots, I. \quad (3)$$

Se si ponesse, per esempio,  $\text{may}(j) \equiv \max\{1, d_j - 2\}$ , si limiterebbe il set di dati utilizzati per la stima agli ultimi tre anni osservati. Questa scelta potrebbe essere motivata dal fatto che l’ipotesi di indipendenza del costo medio dall’AY è ritenuta accettabile solo relativamente ai tre anni di calendario più recenti.

---

<sup>4</sup>Nelle applicazioni dell’FL spesso si effettua, sulla base di un criterio convenzionalmente fissato, una suddivisione preliminare dei sinistri in sinistri *normali* e sinistri *ingenti*, e si tratta separatamente lo sviluppo dei costi ingenti. Questa eventuale suddivisione non verrà qui considerata.

### 3.4 Il modello stocastico per i costi medi

L'uso dello stimatore (3) si giustifica con opportune ipotesi probabilistiche. Si adotterà il seguente modello “time series”.

#### 3.4.1 Ipotesi time series sui i costi medi

**K1.** Gli importi pagati  $C_{i,j}$  di accident year diversi sono indipendenti.

**K2.** Per  $1 \leq i, j \leq I$  esistono costanti  $\kappa_j, \rho_j > 0$  tali che:

$$C_{i,j} = \kappa_j N_{i,j} + \rho_j \sqrt{N_{i,j}} \varepsilon_{i,j}^C, \quad (4)$$

con  $\varepsilon_{i,j}^C$  variabili aleatorie a media 0 e varianza 1, mutuamente indipendenti e indipendenti da  $N_{i,j}$ . ■

In  $t = I$ , fissato  $j$ , la (4) individua per i costi di rimborso una equazione di regressione lineare sui dati  $\{N_{i,j} \mid i \in \mathcal{I}_j\}$ . Una volta scelta un'opportuna forma per la distribuzione di probabilità dei termini di errore  $\varepsilon_{i,j}^C$ , la (4) consente di ricavare l'intera distribuzione di probabilità delle  $C_{i,j}$  per simulazione. Una scelta praticabile si ha assumendo le  $\varepsilon_{i,j}^C$  come v.a. uniformi standard, oppure normali standard.

Sulla base della definizione (2) la (4) si può scrivere, equivalentemente, nella forma:

$$K_{i,j} = \kappa_j + \frac{\rho_j}{\sqrt{N_{i,j}}} \varepsilon_{i,j}^C, \quad (5)$$

interpretabile, per  $j$  fissato, come una regressione lineare pesata (con pesi  $N_{i,j}$ ) con sola intercetta per i costi medi unitari.

#### 3.4.2 Gli stimatori dei parametri del modello stocastico dei costi medi

Dalla teoria classica della regressione lineare si ricava che lo stimatore di  $\kappa_j$  dato dalla (3):

$$\hat{\kappa}_j = \sum_{i \in \mathcal{I}_j} \frac{N_{i,j}}{\sum_{h \in \mathcal{I}_j} N_{h,j}} K_{i,j} = \hat{\kappa}_j^{\text{FL}}, \quad 1 \leq j \leq I. \quad (6)$$

è condizionatamente non distorto dato  $\mathcal{B}_j^N$ ; si ha cioè  $\mathbf{E}(\hat{\kappa}_j | \mathcal{B}_j^N) = \kappa_j$ . Lo stimatore  $\hat{\kappa}_j$  ha varianza (condizionata) minima tra tutti gli stimatori lineari di  $\kappa_j$  e risulta:

$$\mathbf{Var}(\hat{\kappa}_j | \mathcal{B}_j^N) = \frac{\rho_j^2}{\sum_{i \in \mathcal{I}_j} N_{i,j}}, \quad 1 \leq j \leq I. \quad (7)$$

Inoltre, lo stimatore di  $\rho_j^2$ :

$$\hat{\rho}_j^2 = \frac{1}{d_j - \text{may}(j)} \sum_{i \in \mathcal{I}_j} N_{i,j} (K_{i,j} - \hat{\kappa}_j)^2, \quad \text{per } j \text{ tale che } d_j > \text{may}(j), \quad (8)$$

è anch'esso condizionatamente non distorto dato  $\mathcal{B}_j^N$ . Nei casi in cui è  $d_j = \text{may}(j)$  (una sola osservazione) lo stimatore  $\hat{\rho}_j^2$  non è definito e viene ricavato, di solito, per inter/estrapolazione dai valori contigui. In particolare, per dati completi, cioè se è  $\text{may}(j) \equiv 1$ , la stima  $\hat{\rho}_j^2$  si può ricavare come in [6].



### 3.4.3 Una versione distribution free

Gli stimatori (6) e (8), e le relative proprietà, possono essere ottenuti adottando un set di ipotesi più debole delle ipotesi time series K1-K2. Precisamente, si tratta di assumere:

**DFK1.** Gli importi pagati  $C_{i,j}$  di accident year diversi sono indipendenti.

**DFK2.** Per  $1 \leq i, j \leq I$  esistono costanti  $\kappa_j, \rho_j > 0$  tali che:

$$\begin{aligned} \mathbf{E}(C_{i,j} | N_{i,j}) &= \kappa_j N_{i,j}, \\ \mathbf{Var}(C_{i,j} | N_{i,j}) &= \rho_j^2 N_{i,j}. \end{aligned} \tag{9}$$

■

Queste ipotesi su medie e varianze condizionate sono adeguate in una logica “distribution free”, analoga a quella che ha ispirato l’ormai classico modello DFCL di Mack [6]. Siamo qui interessati però a ricavare un modello più flessibile, che consenta di includere facilmente fenomeni aggiuntivi, come per esempio gli effetti diagonali che caratterizzano l’attualizzazione e l’inflazione futura. La via più naturale per ottenere questi risultati è quella simulativa, applicata a partire da una specificata distribuzione di probabilità. Le ipotesi time series K1-K2 consentono agevolmente di realizzare questo obiettivo.

*Osservazione.* La condizione sulla varianza nella (9), scritta nella forma per i costi medi  $\mathbf{Var}(K_{i,j} | N_{i,j}) = \rho_j^2 / N_{i,j}$ , è nota nella teoria della credibilità come condizione di Bühlmann-Straub. In questo contesto  $N_{i,j}$  gioca il ruolo di misura di volume (condizionata).

■

### 3.5 Un modello alternativo

Il fatto che nella (4) la varianza degli importi pagati dipenda dal numero di pagati è il motivo principale che impedisce di assumere l’indipendenza stocastica tra  $C_{i,j}$  e  $N_{i,j}$ . Si potrebbe considerare un modello alternativo per i costi medi sostituendo la **K2** con la seguente ipotesi a varianza quadratica:

**K2’.** Per  $1 \leq i, j \leq I$  esistono costanti  $\kappa_j, \rho_j > 0$  tali che:

$$C_{i,j} = \kappa_j N_{i,j} + \rho_j N_{i,j} \varepsilon_{i,j}^C, \tag{10}$$

con  $\varepsilon_{i,j}^C$  a media 0 e varianza 1, mutuamente indipendenti e indipendenti da  $N_{i,j}$ .

■

Naturalmente la (10) equivale all’ipotesi omoschedastica per i costi medi unitari:

$$K_{i,j} = \kappa_j + \rho_j \varepsilon_{i,j}^C; \tag{11}$$

l’adozione della (11) consente quindi di assumere l’indipendenza stocastica tra  $K_{i,j}$  e  $N_{i,j}$ , con conseguente semplificazione della struttura matematica del modello. Naturalmente, però, gli stimatori coerenti con l’ipotesi **K2’** hanno forma diversa. Si ha:

$$\begin{aligned} \hat{\kappa}_j' &= \frac{1}{d_j - \text{may}(j) + 1} \sum_{i \in \mathcal{I}_j} K_{i,j}, \quad 1 \leq j \leq I, \\ \tilde{\rho}_j^2 &= \frac{1}{d_j - \text{may}(j)} \sum_{i \in \mathcal{I}_j} (K_{i,j} - \hat{\kappa}_j')^2, \quad \text{per } j \text{ tale che } d_j > \text{may}(j). \end{aligned} \tag{12}$$

Lo stimatore dato dalla (12), che esprime  $\hat{\kappa}_j'$  come media aritmetica equipesata dei costi medi osservati, risulta meno frequentemente adottato nelle applicazioni del modello FL deterministico. Per questo motivo si adotterà nel seguito l’ipotesi eteroschedastica data dalla (4) o, equivalentemente, dalla (5).

## 4 Il modello per i numeri

### 4.1 Schemi logici per lo sviluppo dei numeri

Nel modello FL gioca un ruolo rilevante il processo di sviluppo dei numeri, che, come si è detto, è ispirato a una logica di *case outstanding development*: lo sviluppo del numero dei sinistri chiusi  $N_{i,j}$  è “guidato” dallo sviluppo del numero di sinistri messi a riserva  $R_{i,j}$ . Per cui una proiezione del numero degli  $N_{i,j}$  coinvolgerà necessariamente anche una corrispondente proiezione degli  $R_{i,j}$ . In effetti gli andamenti nel tempo di queste due quantità sono tra loro interconnessi, in un processo in cui si alternano decisioni di liquidazione e arrivo di nuova informazione. Si può osservare che i dati di inventario contengono per loro natura informazioni “puntuali” sugli impegni in-essere a livello di singolo sinistro. Gli approcci di tipo *case outstanding development* cercano di utilizzare questa informazione, che qui risulta contenuta, sia pure a livello aggregato, nel valore degli  $R_{i,j}$ .

Conviene tentare una schematizzazione preliminare delle relazioni logiche che governano il processo dello sviluppo interconnesso dei numeri dei chiusi e dei riservati. Si possono considerare dapprima dei “passi elementari”, e poi comporre questi passi in alcuni possibili schemi complessivi di sviluppo.

#### 4.1.1 Passi elementari dello sviluppo

##### I riservati dai riservati

Innanzitutto, per  $1 \leq i \leq I$  e  $\text{mdy}(i) + 1 \leq j \leq I$  vale per il numero dei riservati la relazione ricorrente fondamentale:

$$R_{i,j} = R_{i,j-1} + D_{i,j}^{(1)} + D_{i,j}^{(2)} - D_{i,j}^{(3)} - N_{i,j}. \quad (13)$$

Il numero  $R_{i,j}^{\text{in}}$  di sinistri a riserva all’inizio dell’anno  $(i, j)$  è posto uguale al numero di sinistri a riserva alla fine dell’anno precedente:  $R_{i,j+1}^{\text{in}} = R_{i,j}$ <sup>5</sup>. Si può anche dire che si passa da  $R_{i,j-1}$  a  $R_{i,j}$  applicando alla riserva iniziale il *fattore di sviluppo* nell’anno  $(i, j)$ :

$$F_{i,j} := \frac{R_{i,j}}{R_{i,j-1}},$$

assumendo, ovviamente, il denominatore diverso da zero. Questa relazione di passaggio, che possiamo etichettare “*R-da-R*”, può essere scomposta utilmente in passi intermedi.

##### I con-seguito dai riservati

Si definisca, per  $1 \leq i \leq I$  e  $\text{mdy}(i) + 1 \leq j \leq I$ :

· l’*aggiustamento* (del numero dei riservati):  $D_{i,j} := D_{i,j}^{(1)} + D_{i,j}^{(2)} - D_{i,j}^{(3)}$ ,

e:

· il numero dei *sinistri con seguito*<sup>6</sup>:  $B_{i,j} := R_{i,j-1} + D_{i,j}$ .

Si può quindi considerare una fase di passaggio “*B-da-R*”, in cui nell’anno  $(i, j)$  si passa dai riservati iniziali ai “con-seguito” applicando il termine di aggiustamento  $D_{i,j}$  (sommando cioè i nuovi denunciati e i riaperti e sottraendo i chiusi senza seguito). Si può anche dire che si

<sup>5</sup>Di solito si estende la definizione al primo DY ponendo anche  $R_{i,1}^{\text{in}} := R_{i,0} = D_{i,1}^{(1)}$  e  $D_{i,1}^{(2)} = D_{i,1}^{(3)} = 0$ . Queste convenzioni non sono però essenziali per il seguito.

<sup>6</sup>Anche in questo caso si può estendere la definizione al primo DY assumendo la convenzione  $B_{i,1} := R_{i,0} = D_{i,1}^{(1)}$ .

passa da  $R_{i,j-1}$  a  $B_{i,j}$  applicando alla riserva iniziale il *fattore di aggiustamento* nell'anno  $(i, j)$ :

$$A_{i,j} := \frac{B_{i,j}}{R_{i,j-1}} = 1 + \frac{D_{i,j}}{R_{i,j-1}}.$$

### I chiusi dai con-seguito

I sinistri con seguito costituiscono la “base” dei sinistri ancora da pagare nell'anno  $(i, j)$ , dopo l'arrivo dell'informazione espressa da  $D_{i,j}$ . A seguito dell'attività di liquidazione, verrà pagata in  $(i, j)$  una parte  $N_{i,j}$  di questa base-sinistri e verrà messa a riserva la parte rimanente. Si definisce il *fattore* (o anche *velocità*, o *tasso*) di *liquidazione* nell'anno  $(i, j)$  come:

$$Q_{i,j} := \frac{N_{i,j}}{B_{i,j}};$$

cioè come la frazione liquidata di sinistri con seguito. Si può allora considerare un'altra fase di passaggio “ $N$ -da- $B$ ”, in cui si passa dal numero dei con-seguito al numero dei pagati chiusi applicando il fattore  $Q_{i,j}$ .

Naturalmente varrà sempre la *condizione di chiusura*:

$$R_{i,j} = B_{i,j} - N_{i,j}. \quad (14)$$

### I chiusi dai riservati

Si può considerare anche una fase di passaggio “ $N$ -da- $R$ ”, in cui dai riservati iniziali direttamente si passa ai chiusi applicando il fattore di sviluppo:

$$\Phi_{i,j} := \frac{N_{i,j}}{R_{i,j-1}} = A_{i,j} Q_{i,j}.$$

#### 4.1.2 Schemi di sviluppo dei riservati

Componendo i passi descritti, è possibile definire degli schemi alternativi per la costruzione di modelli di sviluppo per i numeri dei chiusi e dei riservati.

##### Lo schema “in serie”

Un possibile schema di sviluppo si ha osservando che il passaggio  $R$ -da- $R$ , specificato dalla relazione ricorrente (13), può essere costruito effettuando prima il passo  $B$ -da- $R$ , poi il passo  $N$ -da- $B$ , e applicando infine la condizione di chiusura. Con queste definizioni, la (13) si può scrivere come:

$$R_{i,j} = R_{i,j-1} A_{i,j} (1 - Q_{i,j}),$$

che equivale alla fattorizzazione;

$$F_{i,j} = A_{i,j} (1 - Q_{i,j}).$$

Questa sequenza di sviluppo si può schematizzare come in figura 1.

L'approccio può essere definito come schema ricorrente “in serie”: il passo  $R$ -da- $R$  si ottiene componendo (concatenando) il passo  $B$ -da- $R$  e il passo  $N$ -da- $B$ , e ricavando quindi  $R_{i,j}$  come differenza  $B_{i,j} - N_{i,j}$ .

Lo schema in serie può essere espresso in termini di equazioni ricorrenti. Basta imporre la condizione iniziale  $R_{i, \text{mdy}(i)-1} := D_{i, \text{mdy}(i)}$  e richiedere che, per  $1 \leq i \leq I$  e  $\text{mdy}(i) \leq j \leq I$  valga il sistema di equazioni:

$$\begin{cases} B_{i,j} = A_{i,j} R_{i,j-1}, \\ N_{i,j} = Q_{i,j} B_{i,j}, \\ R_{i,j} = B_{i,j} - N_{i,j}. \end{cases}$$

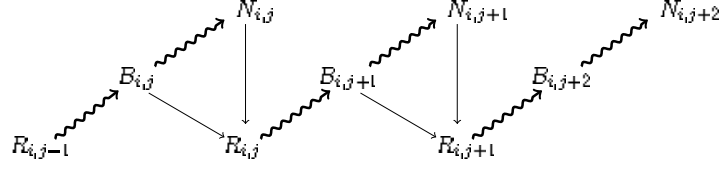


Figura 1: Lo schema ricorrente “in serie”

Evidentemente, la prima equazione corrisponde al passo  $B$ -da- $R$ , la seconda realizza il passo  $N$ -da- $B$  e la terza applica la condizione di chiusura.

### Lo schema “in parallelo”

Se, invece del passo  $N$ -da- $B$ , si utilizza il passo  $N$ -da- $R$  si può scomporre la (13) secondo uno schema che si può definire “in parallelo” (figura 2). Qui il passo  $R$ -da- $R$  si ottiene

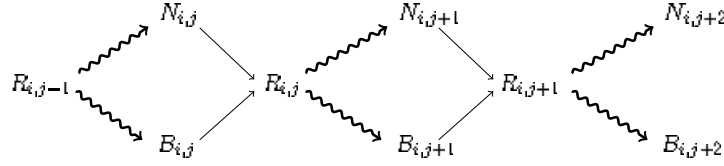


Figura 2: Lo schema ricorrente “in parallelo”

“affiancando”  $B$ -da- $R$  e  $N$ -da- $R$ , e applicando poi ancora la condizione di chiusura. In termini di equazioni ricorrenti, per  $1 \leq i \leq I$  e  $\text{mdy}(i) \leq j \leq I$ , valgono le relazioni:

$$\begin{cases} B_{i,j} = A_{i,j} R_{i,j-1}, \\ N_{i,j} = \Phi_{i,j} R_{i,j-1}, \\ R_{i,j} = B_{i,j} - N_{i,j}, \end{cases}$$

sempre a partire dalla condizione iniziale  $R_{i,\text{mdy}(i)-1} := D_{i,\text{mdy}(i)}$ . In questo caso la seconda equazione realizza il passo  $N$ -da- $R$ , in sostituzione del passo  $N$ -da- $B$ .

Con queste definizioni, la (13) si può anche scrivere come:

$$R_{i,j} = R_{i,j-1} (A_{i,j} - \Phi_{i,j}), \quad (15)$$

che equivale alla rappresentazione:

$$F_{i,j} = A_{i,j} - \Phi_{i,j}.$$

### Lo schema “a catena”

Un ulteriore schema di sviluppo, anch’esso utile nel seguito, si ha considerando una “catena” di passi  $R$ -da- $R$  consecutivi e una serie di passi  $B$ -da- $R$  che si diramano da ciascun “anello”

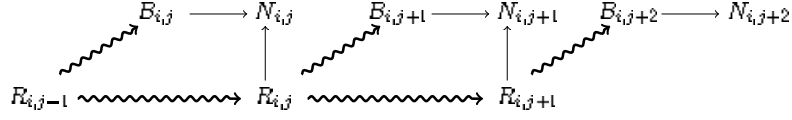


Figura 3: Lo schema ricorrente “a catena”

della catena. Questo schema ha la forma riportata in figura 4, e corrisponde al sistema di equazioni ricorrenti, per  $1 \leq i \leq I$  e  $\text{mdy}(i) \leq j \leq I$ :

$$\begin{cases} R_{i,j} = F_{i,j} R_{i,j-1}, \\ B_{i,j} = A_{i,j} R_{i,j-1}, \\ N_{i,j} = B_{i,j} - R_{i,j}, \end{cases}$$

ancora con la condizione iniziale  $R_{i,\text{mdy}(i)-1} := D_{i,\text{mdy}(i)}$ . Come si vede, qui la condizione di chiusura non è utilizzata per ricavare  $R$  da  $B$  e  $N$ , ma per ricavare  $N$  da  $R$  e  $B$ .

Naturalmente, dato che lo schema in parallelo implica una simmetria tra il processo  $N$  e il processo  $B$ , lo schema a catena si può anche porre nella forma: che corrisponde al sistema,

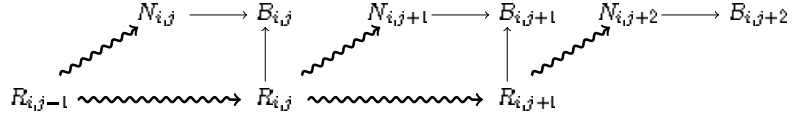


Figura 4: Forma alternativa dello schema a catena

per  $1 \leq i \leq I$  e  $\text{mdy}(i) \leq j \leq I$ :

$$\begin{cases} R_{i,j} = F_{i,j} R_{i,j-1}, \\ N_{i,j} = \Phi_{i,j} R_{i,j-1}, \\ B_{i,j} = N_{i,j} + R_{i,j}. \end{cases}$$

Questa sarà la forma effettivamente utilizzata nel seguito.

#### 4.1.3 La condizione di run-off

Avendo adottato l'ipotesi che lo sviluppo dei rimborsi si completi entro il DY  $I$ , dovrà valere la *condizione di run-off*:

$$R_{i,I} = 0, \quad i = 1, \dots, I; \quad (16)$$

il che implica che per ogni  $i$  debbano valere le uguaglianze:

$$N_{i,I} = B_{i,I}, \quad Q_{i,I} = 1, \quad \Phi_{i,I} = A_{i,I}, \quad F_{i,I} = 0.$$

Si osservi che la (16) dovrebbe essere rispettata sia dai dati, sia dai valori di  $R_{i,I}$  ottenuti “proiettando” lo sviluppo futuro, cioè (visto che si è posto  $J = I$ ) sia per  $i = 1$  che per  $i > 1$ . Tuttavia nelle applicazioni pratiche l'ipotesi di run-off entro il DY  $I$  è spesso intesa come una approssimazione, dato che la condizione di run-off non è effettivamente soddisfatta dai dati. Alcune considerazioni su questo punto verranno sviluppate nel successivo paragrafo 4.2.3.

## 4.2 Il modello deterministico

### 4.2.1 Gli schemi di sviluppo

Il modello FL deterministico può essere descritto adottando indifferentemente sia uno schema di sviluppo “in serie”, sia uno schema “in parallelo”. Si possono quindi adottare le definizioni dei fattori di passaggio già definiti nel paragrafo 4.1, applicando lo schema di sviluppo a partire dai numeri riservati del primo DY  $R_{i,1}$ , che sono comunque sempre osservati. Si assumerà rispettata la condizione di run-off.

**Schema in serie.** Nell’FL impostato secondo lo schema in serie si suppone che per  $1 \leq i \leq I$  e  $\text{mdy}(i) + 1 \leq j \leq I$ , valgano le relazioni di proporzionalità<sup>7</sup>:

$$B_{i,j} = \alpha_j R_{i,j-1}, \quad N_{i,j} = v_j B_{i,j}, \quad (17)$$

dove  $\alpha_j$  e  $v_j$  sono parametri da stimare, con  $v_I = 1$  per la condizione di run-off. Si suppone quindi che il rapporto  $B_{i,j}/R_{i,j-1}$  e il rapporto  $N_{i,j}/B_{i,j-1}$  sia espresso da un parametro  $\alpha_j$  e  $v_j$ , rispettivamente, caratteristico del DY e indipendente dall’AY.

Si ha quindi lo schema ricorrente:

$$\begin{cases} B_{i,j} = \alpha_j R_{i,j-1}, \\ N_{i,j} = v_j B_{i,j}, \\ R_{i,j} = B_{i,j} - N_{i,j}. \end{cases} \quad (18)$$

Evidentemente, la prima equazione corrisponde alla fase *B-da-R*, la seconda realizza la fase *N-da-B* e la terza applica la condizione di chiusura.

**Schema in parallelo.** Nell’FL impostato secondo lo schema in parallelo si suppone che per  $1 \leq i \leq I$  e  $\text{mdy}(i) + 1 \leq j \leq I$ , valga la relazione di proporzionalità:

$$N_{i,j} = \varphi_j R_{i,j-1},$$

con  $\varphi_j$  parametri da stimare. Quindi il rapporto  $N_{i,j}/R_{i,j-1}$  è espresso da un parametro  $\varphi_j$  indipendente dall’AY. Si ha allora lo schema ricorrente:

$$\begin{cases} B_{i,j} = \alpha_j R_{i,j-1}, \\ N_{i,j} = \varphi_j R_{i,j-1}, \\ R_{i,j} = B_{i,j} - N_{i,j}. \end{cases} \quad (19)$$

La seconda equazione realizza la fase *N-da-R*, invece che *N-da-B*.

### 4.2.2 Gli stimatori classici per i numeri

Alla data  $t = I$ , e quindi dato  $\mathcal{D}_I$ , nella pratica attuariale tipicamente si adottano, per  $2 \leq j \leq I$ , gli stimatori:

$$\begin{aligned} \hat{\alpha}_j^{\text{FL}} &:= \sum_{i \in \mathcal{I}_j} \frac{R_{i,j-1}}{\sum_{k \in \mathcal{I}_j} R_{k,j-1}} A_{i,j}, \\ \hat{v}_j^{\text{FL}} &:= \sum_{i \in \mathcal{I}_j} \frac{B_{i,j}}{\sum_{k \in \mathcal{I}_j} B_{k,j}} Q_{i,j}, \end{aligned}$$

---

<sup>7</sup>Distaccandosi dalla notazione più diffusa, converrà definire tutti gli *age-to-age factor*, cioè i parametri di collegamento tra un DY e il successivo, come “ritardati”, cioè con l’indice  $j$  riferito al DY finale. In questo caso, per esempio, sarebbe più usuale definire implicitamente le  $\alpha$  con la relazione  $B_{i,j} = \alpha_{j-1} R_{i,j-1}$ . Stessa osservazione vale per i parametri  $\varphi$  definiti nel seguito.

e:

$$\hat{\varphi}_j^{\text{FL}} := \sum_{i \in \mathcal{I}_j} \frac{R_{i,j-1}}{\sum_{k \in \mathcal{I}_j} R_{k,j-1}} \Phi_{i,j}. \quad (20)$$

Si verifica immediatamente che, per  $2 \leq j \leq I$ , vale la proprietà:

$$\hat{\varphi}_j^{\text{FL}} = \hat{\alpha}_j^{\text{FL}} \hat{v}_j^{\text{FL}}. \quad (21)$$

Se si adotta uno schema in serie, le stime dei numeri di pagati futuri si ricavano quindi sulla base dello schema (18), ponendo, per  $i = 2, \dots, I$  e per  $j \geq d_i$ :

$$\begin{cases} \hat{B}_{i,j}^{\text{FL}} = \hat{\alpha}_j^{\text{FL}} \hat{R}_{i,j-1}^{\text{FL}}, \\ \hat{N}_{i,j}^{\text{FL}} = \hat{v}_j^{\text{FL}} \hat{B}_{i,j}^{\text{FL}}, \\ \hat{R}_{i,j}^{\text{FL}} = \hat{B}_{i,j}^{\text{FL}} - \hat{N}_{i,j}^{\text{FL}}, \end{cases} \quad (22)$$

con  $\hat{R}_{i,d_i}^{\text{FL}} = R_{i,d_i}$ . Dalle (22) si ricava l'espressione esplicita, per  $2 \leq i \leq I$  e  $1 \leq k \leq I - d_i$ :

$$\hat{N}_{i,d_i+k}^{\text{FL}} = R_{i,d_i} \hat{\alpha}_{d_i+k}^{\text{FL}} \hat{v}_{d_i+k}^{\text{FL}} \prod_{h=1}^{k-1} \hat{\alpha}_{d_i+h}^{\text{FL}} (1 - \hat{v}_{d_i+h}^{\text{FL}}) \quad (23)$$

(dove si intende di assegnare valore 1 alle produttorie vuote).

Ci si può basare anche sullo schema in parallelo (19) ponendo, per  $2 \leq i \leq I$  e per  $j \geq d_i$ :

$$\begin{cases} \hat{B}_{i,j}^{\text{FL}} = \hat{\alpha}_j^{\text{FL}} \hat{R}_{i,j-1}^{\text{FL}}, \\ \hat{N}_{i,j}^{\text{FL}} = \hat{\varphi}_j^{\text{FL}} \hat{R}_{i,j-1}^{\text{FL}}, \\ \hat{R}_{i,j}^{\text{FL}} = \hat{B}_{i,j}^{\text{FL}} - \hat{N}_{i,j}^{\text{FL}}, \end{cases}$$

con  $\hat{R}_{i,d_i}^{\text{FL}} = R_{i,d_i}$ . In questo caso si ha l'espressione esplicita, per  $2 \leq i \leq I$  e  $1 \leq k \leq I - d_i$ :

$$\hat{N}_{i,d_i+k}^{\text{FL}} = R_{i,d_i} \hat{\varphi}_{d_i+k}^{\text{FL}} \prod_{h=1}^{k-1} (\hat{\alpha}_{d_i+h}^{\text{FL}} - \hat{\varphi}_{d_i+h}^{\text{FL}}). \quad (24)$$

È rilevante il fatto che, data la (21), le espressioni (23) e (24) sono equivalenti.

Se inoltre si definisce lo stimatore:

$$\hat{f}_j^{\text{FL}} := \sum_{i \in \mathcal{I}_j} \frac{R_{i,j-1}}{\sum_{k \in \mathcal{I}_j} R_{k,j-1}} F_{i,j},$$

è immediato verificare la proprietà, per  $2 \leq j \leq I$ :

$$\hat{f}_j^{\text{FL}} = \hat{\alpha}_j^{\text{FL}} - \hat{\varphi}_j^{\text{FL}} = \hat{\alpha}_j^{\text{FL}} (1 - \hat{v}_j^{\text{FL}}),$$

per cui si ha un'ulteriore espressione equivalente per il numero di chiusi:

$$\hat{N}_{i,d_i+k}^{\text{FL}} = R_{i,d_i} \hat{\varphi}_{d_i+k}^{\text{FL}} \prod_{h=1}^{k-1} \hat{f}_{d_i+h}^{\text{FL}}. \quad (25)$$

Questa espressione equivale ad adottare lo schema ricorrente:

$$\begin{cases} \hat{R}_{i,j}^{\text{FL}} = \hat{f}_j^{\text{FL}} \hat{R}_{i,j-1}^{\text{FL}}, \\ \hat{N}_{i,j}^{\text{FL}} = \hat{\varphi}_j^{\text{FL}} \hat{R}_{i,j-1}^{\text{FL}}, \\ \hat{B}_{i,j}^{\text{FL}} = \hat{N}_{i,j}^{\text{FL}} + \hat{R}_{i,j}^{\text{FL}}, \end{cases}$$

con  $\widehat{R}_{i,d_i}^{\text{FL}} = R_{i,d_i}$ , che è ispirato allo schema a catena già considerato nel paragrafo 4.1.

Analogamente a quanto fatto per i costi medi, si tratta anche qui di costruire un'estensione stocastica coerente dell'approccio deterministico per i numeri. Il punto cruciale è che in ambito stocastico l'equivalenza tra i tre schemi considerati non è più valida. Nonostante lo schema in serie appaia più elegante (e consenta forse una descrizione più “articolata” del meccanismo di formazione dell'incertezza), nel seguito l'estensione stocastica dell'FL verrà sviluppata adottando lo schema in parallelo. I motivi della scelta sono essenzialmente di trattabilità matematica. Come si vedrà, infatti, lo schema a catena consente agevoli soluzioni in forma chiusa, ispirate a quelle dei tipici modelli age-to-age; una caratteristica vantaggiosa dello schema in parallelo è di poter essere collegato allo schema a catena tramite una semplice relazione (c'è una relazione lineare tra i parametri  $f_j$ ,  $\alpha_j$  e  $\varphi_j$ ), cosa che consente di ricavare facilmente espressioni esplicite per i momenti di secondo ordine delle quantità rilevanti.

#### 4.2.3 Includere le code nelle stime

Se l'indice  $I$  rappresentasse effettivamente il run-off dello sviluppo dei numeri di sinistri chiusi si avrebbe  $\widehat{R}_{i,I}^{\text{FL}} = 0$  per ogni  $i$ ; sarebbe cioè rispettata la condizione di run-off (16). Se però la (16) non è soddisfatta dai dati, cioè se è  $R_{1,I} > 0$ , dovrà essere  $B_{1,I} > N_{1,I}$  per cui si avrà:

$$\widehat{v}_{i,I}^{\text{FL}} = \frac{N_{1,I}}{B_{1,I}} < 1.$$

Di conseguenza, per la seconda delle (17) sarà  $\widehat{N}_{i,I}^{\text{FL}} < \widehat{B}_{i,I}^{\text{FL}}$  e quindi  $\widehat{R}_{i,I}^{\text{FL}} > 0$  per  $j > 1$ ; la stima dei sinistri riservati nell'“ultimo” DY risulta cioè positiva. Si pone pertanto il problema di estendere opportunamente lo schema di sviluppo in modo da includere queste “code” nella proiezione.

L'approccio più semplice al problema delle code è quello di introdurre una modifica dei dati, ottenuta annullando i riservati  $R_{1,I}$ , sommando  $R_{1,I}$  ai chiusi  $N_{1,I}$  e sostituendo gli importi  $C_{1,I}$  con  $C_{1,I} (N_{1,I} + R_{1,I}) / N_{1,I}$ . Ciò equivale a “simulare” la liquidazione massiva nel DY  $I$  di tutti i sinistri aperti, a un costo unitario uguale al costo medio stimato  $\widehat{\kappa}_I$ . In ambito stocastico, è rilevante il fatto che questa modifica dei dati avrà effetto sia sulla stima delle OLL, sia sulla stima della corrispondente variabilità. Il metodo può essere considerato analogo a un approccio spesso usato nelle applicazioni del chain ladder stocastico, in cui si somma all'importo pagato  $C_{1,I}$  l'importo residuo  $R_{1,I+1}$  messo a riserva.

*Osservazione.* Un possibile approccio alternativo consiste nel considerare un DY aggiuntivo, l'“anno di sviluppo coda” (*tail development year*)  $I + 1$ . In questo caso si ipotizza che i sinistri residui vengano tutti liquidati nell'anno coda e che in questo anno non vi siano ulteriori aggiustamenti. Si tratta quindi di assumere:

$$\widehat{N}_{i,I+1}^{\text{FL}} = \widehat{B}_{i,I+1}^{\text{FL}} = \widehat{R}_{i,I}^{\text{FL}},$$

il che garantisce  $\widehat{R}_{i,I+1}^{\text{FL}} = 0$  e  $\widehat{N}_{i,j}^{\text{FL}} = 0$  per  $j > I + 1$ . In termini di parametri, l'ipotesi  $\widehat{v}_{i,I}^{\text{FL}} = 1$  va abbandonata, e sostituita con  $\widehat{\alpha}_{i,I+1}^{\text{FL}} = \widehat{v}_{i,I+1}^{\text{FL}} = 1$ , cioè  $\widehat{\varphi}_{i,I+1}^{\text{FL}} = 1$ . Naturalmente alla stima dei numeri coda  $\widehat{N}_{i,I+1}^{\text{FL}}$  andrà affiancata una corrispondente stima del “costo medio coda”  $\widehat{\kappa}_{I+1}^{\text{FL}}$ , in modo da poter ricavare le stime dei “costi coda”  $\widehat{C}_{i,I+1}^{\text{FL}} = \widehat{\kappa}_{I+1}^{\text{FL}} \widehat{N}_{i,I+1}^{\text{FL}}$ . Il punto critico di questo metodo è che in ambito stocastico è anche necessario ricavare una stima dei parametri di varianza per l'anno coda. Una possibilità è quella di adottare un approccio analogo a quello introdotto in [7]; la questione, comunque, non verrà ulteriormente considerata in questo lavoro. ■



### 4.3 Il modello stocastico per i numeri

Come si è detto, l'estensione stocastica dell'FL per la componente numeri verrà effettuata adottando lo schema di sviluppo in parallelo, che utilizza i passi  $N$ -da- $R$  e  $B$ -da- $R$  e la condizione di chiusura.

#### 4.3.1 Il “time series model” per i numeri

Ai fini di avere un modello con distribuzione di probabilità completamente specificata, si adotterà il seguente modello di tipo time series.

##### Ipotesi di modello per i numeri

**N1.** I numeri di riservati  $R_{i,j}$  di accident year diversi sono indipendenti.

**N2.** Per  $1 \leq i \leq I$  e  $2 \leq j \leq I$ , esistono costanti  $\alpha_j$  e  $\beta_j > 0$  tali che:

$$B_{i,j} = \alpha_j R_{i,j-1} + \beta_j \sqrt{R_{i,j-1}} \varepsilon_{i,j}^B, \quad (26)$$

dove le  $\varepsilon_{i,j}^B$  sono v.a. a media 0 e varianza 1, mutuamente indipendenti e indipendenti dalle  $\varepsilon_{i,j}^C$ .

**N3.** Per  $1 \leq i \leq I$  e  $2 \leq j \leq I$ , esistono costanti  $\varphi_j, \psi_j > 0$  e  $\gamma_j \in \mathbb{R}$  tali che:

$$N_{i,j} = \varphi_j R_{i,j-1} + \psi_j \sqrt{R_{i,j-1}} \left( \sqrt{1 - \frac{\gamma_j^2}{\beta_j^2 \psi_j^2}} \varepsilon_{i,j}^N + \frac{\gamma_j}{\beta_j \psi_j} \varepsilon_{i,j}^B \right), \quad (27)$$

dove le  $\varepsilon_{i,j}^N$ , sono v.a. a media 0 e varianza 1, mutuamente indipendenti e indipendenti dalle  $\varepsilon_{i,j}^C$  e dalle  $\varepsilon_{i,j}^B$ , e dove valgono le condizioni:

$$\varphi_I = \alpha_I, \quad \psi_I = \beta_I, \quad \gamma_I = \beta_I^2. \quad (28)$$

■

Le uguaglianze (28), che equivalgono a richiedere l'uguaglianza  $N_{i,I} = B_{i,I}$  per ogni  $i$ , possono essere qualificate come “condizioni no-tail”, nel senso che garantiscono la coerenza del modello previsivo con l'ipotesi di run-off  $J = I$ . Assicurano infatti che l'uguaglianza  $R_{i,I} = 0$  valga con certezza per ogni  $i$ .

*Osservazione.* Come si è detto, se i dati sono coerenti con l'ipotesi no-tail dovrà a aversi  $N_{1,I} = B_{1,I}$ , cioè  $R_{1,I} = 0$ . Data la forma dello stimatore di  $\varphi_I$ , ciò implicherà  $\hat{\varphi}_I = \hat{\alpha}_I$ . In assenza delle (28) però, questa uguaglianza garantirà che la proprietà  $R_{i,I} = 0, 2 \leq i \leq I$ , valga solo in aspettativa. ■

*Osservazione.* Per la condizione di chiusura (14), dato cioè che deve essere  $R_{i,j} = B_{i,j} - N_{i,j}$ , la (26) e la (27) hanno il carattere di equazioni ricorrenti. A rigore però, queste relazioni presentano una forma di incoerenza. Per evitare che nella “serie storica” delle  $R_{i,j}$  si producano valori negativi che renderebbero non definita l'operazione di radice quadrata, sarebbe necessario limitare opportunamente il supporto della distribuzione di probabilità delle  $\varepsilon_{i,j}^N$  e delle  $\varepsilon_{i,j}^B$  in funzione dei valori raggiunti da  $R_{i,j-1}$ ,  $B_{i,j-1}$  e  $N_{i,j-1}$ , cosa che sarebbe però in contraddizione con l'ipotesi di indipendenza stocastica dei termini di errore. Si tratta dello stesso problema di incoerenza rilevato in [8], p. 544, a proposito della versione BMW [1] del modello DFCL di Mack. Tecnicamente, per rimuovere questo problema “di square root” in modo rigoroso sarebbe necessario introdurre una debole (ma complicata!) forma di correlazione tra i termini di errore.

Se si considerano modelli basati su quantità cumulate, quindi con trend crescente al crescere del DY (come nel modello BBMW), il problema di square root non ha rilevanza dal punto di vista della effettiva praticabilità algoritmica e può essere aggirato con semplici accorgimenti ad hoc. La gestione del problema può richiedere maggiori attenzioni nel caso di modelli definiti su quantità incrementali, come quello qui considerato. ■

Date le ipotesi di indipendenza sulle  $\varepsilon_{i,j}^B$  e sulle  $\varepsilon_{i,j}^N$ , si dimostra immediatamente la proprietà:

$$\mathbf{Cov}(B_{i,j}, N_{i,j} | R_{i,j-1}) = \gamma_j R_{i,j-1}, \quad 1 \leq i \leq I, \quad 2 \leq j \leq I. \quad (29)$$

La covarianza condizionata considerata nella (29) non va confusa con la covarianza tra  $B_{i,j}$  e  $N_{i,j}$  dato  $\mathcal{D}_I$ . Si ha, per  $2 \leq i \leq I$  e  $d_i + 1 \leq j \leq I$ :

$$\begin{aligned} \mathbf{Cov}(B_{i,j}, N_{i,j} | \mathcal{D}_I) &= \mathbf{E} \left[ \mathbf{Cov}(B_{i,j}, N_{i,j} | R_{i,j-1}) | \mathcal{D}_I \right] \\ &\quad + \mathbf{Cov} \left[ \mathbf{E}(B_{i,j} | R_{i,j-1}), \mathbf{E}(N_{i,j} | R_{i,j-1}) | \mathcal{D}_I \right]. \end{aligned}$$

Quindi per la (29) e dato che è  $\mathbf{E}(B_{i,j} | R_{i,j-1}) = \alpha_j R_{i,j-1}$  e  $\mathbf{E}(N_{i,j} | R_{i,j-1}) = \varphi_j R_{i,j-1}$ , si ha:

$$\mathbf{Cov}(B_{i,j}, N_{i,j} | \mathcal{D}_I) = \gamma_j \mathbf{E}(R_{i,j-1} | \mathcal{D}_I) + \alpha_j \varphi_j \mathbf{Var}(R_{i,j-1} | \mathcal{D}_I). \quad (30)$$

*Osservazione.* I parametri di covarianza  $\gamma_j$  risulteranno di difficile determinazione, data la scarsità di dati, in particolare per valori di  $j$  prossimi a  $I$ . Nelle applicazioni pratiche potrà essere ragionevole porre  $\gamma_j = 0$ ,  $2 \leq j < I$ , assumendo quindi una ipotesi di non-correlazione condizionata tra i chiusi e i con-seguito. Naturalmente, anche se la covarianza dato  $R_{i,j-1}$  è nulla, la covarianza dato  $\mathcal{D}_I$  è ancora positiva, come si vede dalla (30). ■

#### 4.3.2 Il processo stocastico dei numeri

Alla data  $t = I$ , il processo stocastico relativo ai valori futuri  $\{N_{i,j}, R_{i,j}\} \in \mathcal{F}_I$  resta caratterizzato includendo le relazioni (26) e (27) in uno schema in parallelo. Per  $i = 2, \dots, I$  e per  $j \geq d_i$  si ha lo schema ricorrente stocastico:

$$\begin{cases} B_{i,j} = \alpha_j R_{i,j-1} + \beta_j \sqrt{R_{i,j-1}} \varepsilon_j^B, \\ N_{i,j} = \varphi_j R_{i,j-1} + \psi_j \sqrt{R_{i,j-1}} \left( \varepsilon_{i,j}^N \sqrt{1 - \gamma_j^2 / (\beta_j^2 \psi_j^2)} + \varepsilon_{i,j}^B \gamma_j / (\beta_j \psi_j) \right), \\ R_{i,j} = B_{i,j} - N_{i,j}. \end{cases}$$

#### 4.3.3 Gli stimatori dei parametri del modello stocastico dei numeri

Risulta che, in  $t = I$ :

$$\hat{\alpha}_j = \frac{\sum_{i \in \mathcal{I}_j} B_{i,j}}{\sum_{i \in \mathcal{I}_j} R_{i,j-1}}, \quad 2 \leq j \leq I, \quad (31)$$

e:

$$\hat{\varphi}_j = \frac{\sum_{i \in \mathcal{I}_j} N_{i,j}}{\sum_{i \in \mathcal{I}_j} R_{i,j-1}}, \quad 2 \leq j \leq I, \quad (32)$$

sono stimatori condizionatamente non distorti, dato  $\mathcal{B}_{j-1}^R$ , di  $\alpha_j$  e  $\varphi_j$ , rispettivamente. Questi stimatori sono a varianza minima tra tutti gli stimatori lineari, e si ha, per  $2 \leq j \leq I$ :

$$\mathbf{Var}(\hat{\alpha}_j | \mathcal{B}_{j-1}^R) = \frac{\beta_j^2}{\sum_{i \in \mathcal{I}_j} R_{i,j-1}}, \quad \mathbf{Var}(\hat{\varphi}_j | \mathcal{B}_{j-1}^R) = \frac{\psi_j^2}{\sum_{i \in \mathcal{I}_j} R_{i,j-1}}. \quad (33)$$

Inoltre:

$$\widehat{\beta}_j^2 = \frac{1}{d_j - \text{may}(j)} \sum_{i \in \mathcal{I}_j} R_{i,j-1} \left( \frac{B_{i,j}}{R_{i,j-1}} - \widehat{\alpha}_j \right)^2, \quad (34)$$

$$\widehat{\psi}_j^2 = \frac{1}{d_j - \text{may}(j)} \sum_{i \in \mathcal{I}_j} R_{i,j-1} \left( \frac{N_{i,j}}{R_{i,j-1}} - \widehat{\varphi}_j \right)^2, \quad (35)$$

e:

$$\widehat{\gamma}_j = \frac{1}{d_j - \text{may}(j)} \sum_{i \in \mathcal{I}_j} R_{i,j-1} \left( \frac{B_{i,j}}{R_{i,j-1}} - \widehat{\alpha}_j \right) \left( \frac{N_{i,j}}{R_{i,j-1}} - \widehat{\varphi}_j \right), \quad (36)$$

sono stimatori condizionatamente non distorti di  $\beta_j^2$ ,  $\psi_j^2$  e  $\gamma_j$ , rispettivamente, restando inteso che questi stimatori sono definiti per  $j$  tale che  $d_j > \text{may}(j)$ . Quando tale condizione non è rispettata si procederà ancora per inter/estrapolazione.

Si ha quindi:

$$\widehat{\alpha}_j = \widehat{\alpha}_j^{\text{FL}}, \quad \widehat{\varphi}_j = \widehat{\varphi}_j^{\text{FL}};$$

risulta cioè che gli stimatori dei parametri di “slope” delle regressioni per i numeri coincidono con i corrispondenti stimatori puntuali adottati nell’FL.

#### 4.3.4 La relazione con lo schema a catena

Dalle ipotesi N1-N3, essendo  $R_{i,j} = B_{i,j} - N_{i,j}$ , discende che sia, per  $1 \leq i \leq I$  e  $2 \leq j \leq I$ :

$$R_{i,j} = (\alpha_j - \varphi_j) R_{i,j-1} + \sqrt{R_{i,j-1}} \left[ (\beta_j - \gamma_j / \beta_j) \varepsilon_j^B - \sqrt{\psi_j^2 - \gamma_j^2 / \beta_j^2} \varepsilon_j^N \right]. \quad (37)$$

Ne consegue:

$$\mathbf{E}(R_{i,j} | R_{i,j-1}) = (\alpha_j - \varphi_j) R_{i,j-1}, \quad (38)$$

e, per l’indipendenza:

$$\mathbf{Var}(R_{i,j} | R_{i,j-1}) = R_{i,j-1} (\beta_j^2 - 2\gamma_j + \psi_j^2). \quad (39)$$

Queste proprietà delle medie e delle varianze condizionate possono essere strutturate assumendo che, per  $1 \leq i \leq I$  e  $2 \leq j \leq I$ , si abbia:

$$R_{i,j} = f_j R_{i,j-1} + \sigma_j \sqrt{R_{i,j-1}} \varepsilon_{i,j}^R, \quad (40)$$

dove le  $\varepsilon_{i,j}^R$  sono v.a. a media 0 e varianza 1, mutuamente indipendenti e indipendenti dalle  $\varepsilon_{i,j}^C$ , e avendo posto:

$$f_j := \alpha_j - \varphi_j, \quad \sigma_j^2 := \beta_j^2 - 2\gamma_j + \psi_j^2.$$

Anche se la (40) non è in tutti i sensi una conseguenza delle (26) e (27) e della condizione di chiusura, la (38) e la (39) lo sono; e la (40), che implica la (38) e la (39), può essere utilizzata come relazione comoda per ricavare espressioni di media e varianza di  $R_{i,j} \in \mathcal{F}_I$ , secondo le linee dimostrative usate nei modelli age-to-age.

#### 4.4 Un modello alternativo

Anche per i numeri  $N_{i,j}$  e  $B_{i,j}$  si potrebbe considerare un modello alternativo a quello caratterizzato dalle **N2** e **N3**, sostituendo la (26) e la (27) con le relazioni a varianza quadratica:

$$B_{i,j} = \alpha_j R_{i,j-1} + \beta_j R_{i,j-1} \varepsilon_{i,j}^B, \quad (41)$$

e:

$$N_{i,j} = \varphi_j R_{i,j-1} + \psi_j R_{i,j-1} \left( \sqrt{1 - \frac{\gamma_j^2}{\beta_j^2 \psi_j^2}} \varepsilon_{i,j}^N + \frac{\gamma_j}{\beta_j \psi_j} \varepsilon_{i,j}^B \right), \quad (42)$$

Si eviterebbe in questo modo il problema dell'incongruenza dovuta a eventuali valori negativi di  $R_{i,j}$ . Tuttavia, gli stimatori coerenti con la (41) e la (42) avrebbero minore significatività pratica. In particolare, le stime per  $\alpha_j$  e  $\varphi_j$  sarebbero ottenute come medie equipesate:

$$\hat{\alpha}'_j = \frac{1}{d_j - \text{may}(j) + 1} \sum_{i \in \mathcal{I}_j} \frac{B_{i,j}}{R_{i,j-1}}, \quad \hat{\varphi}'_j = \frac{1}{d_j - \text{may}(j) + 1} \sum_{i \in \mathcal{I}_j} \frac{N_{i,j}}{R_{i,j-1}},$$

senza tener conto delle differenze di “volume” dei diversi accident year.

### 5 Il modello stocastico per gli importi e la stima delle OLL

Il modello time series globale, che caratterizza gli importi pagati  $C_{i,j}$  tramite l'approccio frequency-severity introdotto nel par. 3.1, si ottiene semplicemente adottando sia le ipotesi K1-K2 per i costi medi esposte nel par. 3.4, sia le ipotesi N1-N3 per i numeri esposte nel par. 4.3, strutturate secondo lo schema in parallelo. Il modello che si ottiene è di tipo trivariato; è caratterizzato cioè da tre fonti di incertezza: i termini di errore  $\varepsilon_{i,j}^C$  che generano il processo dei costi medi, i termini di errore  $\varepsilon_{i,j}^B$  relativi al processo del numero dei con-seguito e i termini di errore  $\varepsilon_{i,j}^N$  caratteristici del processo del numero di chiusi; questi termini di errore sono caratterizzati dalle ipotesi distribuzionali e di indipendenza già descritte. Considerato al tempo  $t = I$ , il modello time series trivariato equivale ad adottare il sistema ricorrente stocastico:

$$\begin{cases} B_{i,j} = \alpha_j R_{i,j-1} + \beta_j \sqrt{R_{i,j-1}} \varepsilon_j^B, \\ N_{i,j} = \varphi_j R_{i,j-1} + \psi_j \sqrt{R_{i,j-1}} \left( \varepsilon_{i,j}^N \sqrt{1 - \gamma_j^2 / (\beta_j^2 \psi_j^2)} + \varepsilon_{i,j}^B \gamma_j / (\beta_j \psi_j) \right), \\ R_{i,j} = B_{i,j} - N_{i,j}, \\ C_{i,j} = \kappa_j N_{i,j} + \rho_j \sqrt{N_{i,j}} \varepsilon_j^C, \end{cases} \quad (43)$$

definito per  $2 \leq i \leq I$  e  $j > d_i$  a partire dal numero dei riservati  $R_{i,d_i}$  osservati in  $t = I$ . Gli stimatori per i parametri del modello hanno la forma e le proprietà già precisate. Insomma, il set completo di ipotesi K1-K2 e N1-N3 è sufficiente per generare per simulazione le distribuzioni di probabilità di tutte le variabili che costituiscono gli obiettivi del modello. Tali ipotesi consentono però anche di ricavare utili espressioni in forma chiusa per gli stimatori di medie e varianze. A questo fine, è conveniente formulare l'approccio in parallelo secondo uno schema a catena; ciò equivale a riformulare le (43) nella forma:

$$\begin{cases} R_{i,j} = f_j R_{i,j-1} + \sqrt{R_{i,j-1}} [(\beta_j - \gamma_j / \beta_j) \varepsilon_j^B - \sqrt{\psi_j^2 - \gamma_j^2 / \beta_j^2} \varepsilon_j^N], \\ N_{i,j} = \varphi_j R_{i,j-1} + \sqrt{R_{i,j-1}} [(\gamma_j / \beta_j) \varepsilon_j^B + \sqrt{\psi_j^2 - \gamma_j^2 / \beta_j^2} \varepsilon_j^N], \\ B_{i,j} = N_{i,j} + R_{i,j}, \\ C_{i,j} = \kappa_j N_{i,j} + \rho_j \sqrt{N_{i,j}} \varepsilon_j^C, \end{cases}$$

avendo posto  $f_j = \alpha_j - \varphi_j$ . Ci occuperemo ora degli stimatori per le medie. Le espressioni per gli stimatori delle varianze saranno ricavate nel paragrafo 6.

## 5.1 Gli stimatori per le OLL attese

Il primo obiettivo di qualsiasi modello di loss reserving è quello di fornire una stima del valore atteso degli impegni in essere, e in particolare delle OLL, intese come somma degli importi ancora da sborsare per chiudere tutti i sinistri aperti. In un modello frequency-severity la stima delle OLL in termini di importi si ottiene componendo la stima delle OLL in termini di numeri con la stima dei costi medi. Sarà utile quindi procedere per gradi, ricavando prima gli stimatori per il numero atteso dei riservati, quindi quelli per il numero atteso di chiusi e infine quello per i costi attesi.

### 5.1.1 Numero dei riservati

Si trova, per  $2 \leq i \leq I$  e  $d_i + 1 \leq j \leq I$ , che il valore atteso dato  $\mathcal{D}_I$  di  $R_{i,j}$  è:

$$\mathbf{E}(R_{i,j}|\mathcal{D}_I) = R_{i,d_i} \prod_{k=d_i+1}^j f_k. \quad (44)$$

e che:

$$\widehat{R}_{i,j} := R_{i,d_i} \prod_{k=d_i+1}^j \widehat{f}_k, \quad (45)$$

con  $\widehat{f}_k := \widehat{\alpha}_k - \widehat{\varphi}_k$ , è uno stimatore non distorto di  $\mathbf{E}(R_{i,j}|\mathcal{D}_I)$ .

*Dimostrazione.* La (44) si ottiene con ripetute applicazioni dell'operatore di media condizionata:

$$\begin{aligned} \mathbf{E}(R_{i,j}|\mathcal{D}_I) &= \mathbf{E}[\mathbf{E}(R_{i,j}|R_{i,j-1})|\mathcal{D}_I] \\ &= f_j \mathbf{E}(R_{i,j-1}|\mathcal{D}_I) \\ &= f_j f_{j-1} \mathbf{E}(R_{i,j-2}|\mathcal{D}_I) \\ &\dots \\ &= f_j f_{j-1} \dots f_{d_i+1} R_{i,d_i}. \end{aligned}$$

Applicando analogo procedimento a  $\widehat{R}_{i,j}$  e dato che gli stimatori  $\widehat{f}_j$  sono non distorti si ha:

$$\begin{aligned} \mathbf{E}(\widehat{R}_{i,j}|\mathcal{D}_I) &= \mathbf{E}[\mathbf{E}(\widehat{R}_{i,j}|\mathcal{B}_{j-1}^R)|\mathcal{D}_I] \\ &= f_j \mathbf{E}(\widehat{R}_{i,j-1}|\mathcal{D}_I) \\ &= f_j f_{j-1} \mathbf{E}(\widehat{R}_{i,j-2}|\mathcal{D}_I) \\ &\dots \\ &= f_j f_{j-1} \dots f_{d_i+1} R_{i,d_i}, \end{aligned}$$

cioè  $\mathbf{E}(\widehat{R}_{i,j}|\mathcal{D}_I) = \mathbf{E}(R_{i,j}|\mathcal{D}_I)$ . ■

### 5.1.2 Numero dei pagati chiusi

Per  $2 \leq i \leq I$  e  $d_i + 1 \leq j \leq I$ , il valore atteso dato  $\mathcal{D}_I$  di  $N_{i,j}$  è (intendendo di assegnare il valore 1 alle produttorie vuote):

$$\mathbf{E}(N_{i,j}|\mathcal{D}_I) = R_{i,d_i} \varphi_j \prod_{k=d_i+1}^{j-1} f_k; \quad (46)$$

inoltre:

$$\widehat{N}_{i,j} := R_{i,d_i} \widehat{\varphi}_j \prod_{k=d_i+1}^{j-1} \widehat{f}_k = \widehat{\varphi}_j \widehat{R}_{i,j-1}, \quad (47)$$

è uno stimatore non distorto di  $\mathbf{E}(N_{i,j}|\mathcal{D}_I)$ .

*Dimostrazione.* Infatti si ha  $\mathbf{E}(N_{i,j}|\mathcal{D}_I) = \varphi_j \mathbf{E}(R_{i,j-1}|\mathcal{D}_I)$ . Inoltre:

$$\begin{aligned} \mathbf{E}(\widehat{N}_{i,j}|\mathcal{D}_I) &= \mathbf{E}[\mathbf{E}(\widehat{N}_{i,j}|\mathcal{B}_{j-1}^R)|\mathcal{D}_I] \\ &= \varphi_j \mathbf{E}(\widehat{R}_{i,j-1}|\mathcal{D}_I) \\ &= \varphi_j f_{j-1} \mathbf{E}(\widehat{R}_{i,j-2}|\mathcal{D}_I) \\ &\dots \\ &= \varphi_j f_{j-1} \dots f_{d_i+1} R_{i,d_i}, \end{aligned}$$

cioè  $\mathbf{E}(\widehat{N}_{i,j}|\mathcal{D}_I) = \mathbf{E}(N_{i,j}|\mathcal{D}_I)$ . ■

Data la coincidenza con gli stimatori FL dei parametri, questa espressione equivale alla (23), alla (24) e alla (25).

### 5.1.3 Costi

Per  $2 \leq i \leq I$  e  $d_i + 1 \leq j \leq I$ , il valore atteso dato  $\mathcal{D}_I$  di  $C_{i,j}$  è:

$$\mathbf{E}(C_{i,j}|\mathcal{D}_I) = R_{i,d_i} \kappa_j \varphi_j \prod_{k=d_i+1}^{j-1} f_k, \quad (48)$$

e:

$$\widehat{C}_{i,j} := R_{i,d_i} \widehat{\kappa}_j \widehat{\varphi}_j \prod_{k=d_i+1}^{j-1} \widehat{f}_k = \widehat{\kappa}_j \widehat{N}_{i,j},$$

è uno stimatore non distorto di  $\mathbf{E}(C_{i,j}|\mathcal{D}_I)$ .

*Dimostrazione.* Infatti la (48) si ottiene subito osservando che:

$$\mathbf{E}(C_{i,j}|\mathcal{D}_I) = \mathbf{E}(K_{i,j} N_{i,j}|\mathcal{D}_I) = \mathbf{E}[\mathbf{E}(K_{i,j}|N_{i,j}) \cdot N_{i,j}|\mathcal{D}_I] = \kappa_j \mathbf{E}(N_{i,j}|\mathcal{D}_I).$$

Si ha poi, ricordando che è  $\mathbf{E}(\widehat{\kappa}_j|\mathcal{B}_j^N) = \kappa_j$ :

$$\begin{aligned} \mathbf{E}(\widehat{C}_{i,j}|\mathcal{D}_I) &= \mathbf{E}[\mathbf{E}(\widehat{C}_{i,j}|\mathcal{B}_j^N)|\mathcal{D}_I] \\ &= \kappa_j \mathbf{E}[\mathbf{E}(\widehat{N}_{i,j}|\mathcal{B}_j^N)|\mathcal{D}_I] \\ &= \kappa_j \mathbf{E}(\widehat{N}_{i,j}|\mathcal{D}_I) \\ &= \kappa_j \varphi_j f_{j-1} \dots f_{d_i+1} R_{i,d_i}, \end{aligned}$$

cioè  $\mathbf{E}(\widehat{C}_{i,j}|\mathcal{D}_I) = \mathbf{E}(C_{i,j}|\mathcal{D}_I)$ . ■

Se si prescinde dall'inflazione futura, la relazione  $\widehat{C}_{i,j} = \widehat{\kappa}_j \widehat{N}_{i,j}$ , corredata dalla (6) e dalla (47), riproduce il nucleo essenziale del tradizionale processo di stima puntuale della riserva sinistri col metodo FL.

## 5.2 Una versione distribution free del modello per gli importi

Nel paragrafo 3.4.3 si è già osservato che le ipotesi time series sul processo dei costi medi possono essere riformulate nella forma di ipotesi più deboli, sulle medie e le varianze condizionate, dando origine a una versione del modello di tipo distribution free. Lo stesso ragionamento può essere applicato all'intero modello trivariato, che può essere ridotto a una versione distribution free adottando le seguenti ipotesi su medie e varianze condizionate.

### Ipotesi distribution free

**DF1.** Gli importi pagati  $C_{i,j}$ , i numeri di sinistri chiusi  $N_{i,j}$  e i numeri di sinistri con seguito  $B_{i,j}$  di diversi accident year sono indipendenti.

**DF2.** Per  $1 \leq i \leq I$  esistono costanti  $\kappa_j, \varphi_j, \alpha_j > 0$  tali che:

$$\begin{aligned} \mathbf{E}(C_{i,j}|N_{i,mdy(i)}, \dots, N_{i,j}) &= \kappa_j N_{i,j}, & 1 \leq j \leq I, \\ \mathbf{E}(B_{i,j}|R_{i,mdy(i)}, \dots, R_{i,j-1}) &= \alpha_j R_{i,j-1}, & 2 \leq j \leq I, \\ \mathbf{E}(N_{i,j}|R_{i,mdy(i)}, \dots, R_{i,j-1}) &= \varphi_j R_{i,j-1}, & 2 \leq j \leq I, \end{aligned}$$

con  $\varphi_I = \alpha_I$ .

**DF3.** Per  $1 \leq i \leq I$  esistono costanti  $\rho_j, \psi_j, \beta_j > 0$  tali che:

$$\begin{aligned} \mathbf{Var}(C_{i,j}|N_{i,mdy(i)}, \dots, N_{i,j}) &= \rho_j^2 N_{i,j}, & 1 \leq j \leq I, \\ \mathbf{Var}(B_{i,j}|R_{i,mdy(i)}, \dots, R_{i,j-1}) &= \beta_j^2 R_{i,j-1}, & 2 \leq j \leq I, \\ \mathbf{Var}(N_{i,j}|R_{i,mdy(i)}, \dots, R_{i,j-1}) &= \psi_j^2 R_{i,j-1}, & 2 \leq j \leq I, \end{aligned}$$

con  $\psi_I = \beta_I$ .

**DF4.** Per  $1 \leq i \leq I$  e  $2 \leq j \leq I$ :

$$\mathbf{Cov}(B_{i,j}, N_{i,j}|R_{i,mdy(i)}, \dots, R_{i,j-1}) = \gamma_j R_{i,j-1},$$

con  $\gamma_I = \beta_I^2$ . ■

Si dimostra facilmente che le condizioni DF1-DF4 sono sufficienti affinché valgano tutte le proprietà sia degli stimatori dei parametri descritte nei paragrafi 3.4.2 e 4.3.3, sia degli stimatori di  $R_{i,j}$ ,  $N_{i,j}$  e  $C_{i,j}$ , ricavate nei paragrafi 5.1.1, 5.1.2 e 5.1.3.

## 6 Gli stimatori per il Mean Square Error of Prediction

È ampiamente riconosciuta, in particolare nelle applicazioni in Solvency II, la necessità di includere nella misurazione dell'incertezza anche la componente derivante dagli errori di stima dei parametri alla base dei modelli. Questa esigenza vale, naturalmente, anche per i momenti di secondo ordine delle distribuzioni di probabilità delle variabili rilevanti, e porta a definire un concetto “esteso” di varianza noto come *Mean Square Error of Prediction* (MSEP). In questo paragrafo si ricaveranno espressioni in forma chiusa per gli stimatori del MSEP delle OLL, intese sia in termini di numeri che in termini di importi. Queste espressioni saranno anche utili per tenere sotto controllo l'accuratezza numerica delle più generali procedure di simulazione che saranno illustrate successivamente.

Facendo riferimento, per fissare le idee, ai costi  $C_{i,j}$  per  $i + j > I + 1$ , il MSEP condizionato di  $C_{i,j}$ , dato  $\mathcal{D}_I$ , è espresso dalla:

$$\text{mse}_{C_{i,j}|\mathcal{D}_I}(\widehat{C_{i,j}}) = \mathbf{Var}(C_{i,j}|\mathcal{D}_I) + [\widehat{C_{i,j}} - \mathbf{E}(C_{i,j}|\mathcal{D}_I)]^2. \quad (49)$$

È dato quindi dalla somma della varianza condizionata  $\mathbf{Var}(C_{i,j}|\mathcal{D}_I)$ , che esprime la componente “di processo” (*process variance, process error*), e dello scarto quadratico  $[\widehat{C}_{i,j} - \mathbf{E}(C_{i,j}|\mathcal{D}_I)]^2$ , che rappresenta la componente “di stima” (*estimation variance, estimation error*). Il MSEF a sua volta è anche noto come *prediction variance* o *prediction error*.

Una stima del prediction error richiede quindi di ricavare dai dati  $\mathcal{D}_I$  una stima  $\widehat{\mathbf{Var}}(C_{i,j}|\mathcal{D}_I)$  per il process error, che indicheremo con  $\widehat{\mathbf{Var}}(C_{i,j}|\mathcal{D}_I)$ , e una stima  $\widehat{\mathcal{E}}([\widehat{C}_{i,j} - \mathbf{E}(C_{i,j}|\mathcal{D}_I)]^2|\mathcal{D}_I)$  per l'estimation error.

Il calcolo di grandezze di questo tipo può essere una questione complessa. In particolare, sia che si proceda in forma chiusa sia che si passi per via simulativa, la stima della componente di estimation error richiede l'introduzione di opportune ipotesi definitorie aggiuntive.

## 6.1 Estimation error e conditional resampling

Dato che è  $C_{i,j} = \kappa_j N_{i,j}$ , lo scarto quadratico  $[\widehat{C}_{i,j} - \mathbf{E}(C_{i,j}|\mathcal{D}_I)]^2$  che esprime l'estimation error può essere scritto come:

$$[\widehat{C}_{i,j} - \mathbf{E}(C_{i,j}|\mathcal{D}_I)]^2 = [\widehat{\kappa}_j \widehat{N}_{i,j} - \kappa_j \mathbf{E}(N_{i,j}|\mathcal{D}_I)]^2,$$

dove  $\widehat{N}_{i,j}$  è fornito dalla (47); quindi (dato che è  $\widehat{f}_j = \widehat{\alpha}_j - \widehat{\varphi}_j$ ) la stima  $\widehat{\kappa}_j \widehat{N}_{i,j}$  risulterà espressa come una specificata funzione degli stimatori  $\widehat{\kappa}_j, \widehat{\varphi}_j, \widehat{\alpha}_j$ . Per la (46), anche l'aspettativa  $\kappa_j \mathbf{E}(N_{i,j}|\mathcal{D}_I)$  sarà a sua volta espressa come una specificata funzione dei parametri  $\kappa_j, \varphi_j, \alpha_j$ , che sono incogniti. Per ricavare una stima dello scarto quadratico della stima si tratta quindi, intuitivamente, di ricavare una stima della variabilità dei valori stimati  $\widehat{\kappa}_j, \widehat{\varphi}_j, \widehat{\alpha}_j$  rispetto ai valori “veri”  $\kappa_j, \varphi_j, \alpha_j$ . Queste volatilità vengono di solito stimate immaginando di effettuare un *ricampionamento* (*resampling*) dei dati osservati, con conseguente “nuova stima” dei parametri. Analogamente alla versione del DFCL introdotto in [1], il time series model per i numeri proposto nel paragrafo 4 non è a incrementi indipendenti; ne consegue che esistono vari modi di effettuare il ricampionamento relativo ai parametri  $\widehat{\varphi}_j$  e  $\widehat{\alpha}_j$ , ciascuno dei quali corrisponde a un diverso modo di schematizzare la struttura di incertezza del problema. Una rassegna dei possibili approcci al ricampionamento si può trovare in [10], in particolare, con riferimento al DFCL, nel paragrafo 7.4.

Si utilizzerà qui l'approccio cosiddetto di *ricampionamento condizionato* (*conditional resampling*), analogamente a quanto fatto per il DFCL in [1]. Si indicherà con  $\widehat{\mathbf{E}}_{\mathcal{D}_I}(x)$  lo stimatore di  $x$  dato  $\mathcal{D}_I$ , ottenuto in ipotesi di ricampionamento condizionato.

Per definizioni puntuali e dettagli sul conditional resampling si rimanda a [1]; ci limiteremo qui a illustrare il principio ispiratore dell'approccio. Come si è detto, la logica di misurazione dell'estimation error si basa sulla misurazione della variabilità degli stimatori  $\widehat{\kappa}_j, \widehat{\varphi}_j$  e  $\widehat{\alpha}_j$  (come funzioni dei dati) rispetto ai corrispondenti “valori veri”  $\kappa_j, \varphi_j$  e  $\alpha_j$ . Si richiede quindi di definire la modalità con cui si immagina che i dati vengano ricampionati. Nelle procedure di calcolo il punto cruciale consiste, per un AY  $i$  fissato, nello stabilire come ricampionare prodotti del tipo:

$$\widehat{f}_{d_i+1}^2 \cdot \widehat{f}_{d_i+2}^2 \cdot \dots \cdot \widehat{f}_{j-1}^2 \cdot \widehat{f}_j \widehat{\varphi}_j \widehat{\kappa}_j. \quad d_i < j \leq I.$$

Nel conditional resampling la misurazione della variabilità si ottiene prendendo le aspettative condizionate:

$$\mathbf{E}(\widehat{f}_{d_i+1}^2|\mathcal{B}_{d_i}^R) \cdot \mathbf{E}(\widehat{f}_{d_i+2}^2|\mathcal{B}_{d_i+1}^R) \cdot \dots \cdot \mathbf{E}(\widehat{f}_{j-1}^2|\mathcal{B}_{j-2}^R) \cdot \mathbf{E}(\widehat{f}_j \widehat{\varphi}_j|\mathcal{B}_{j-1}^R) \cdot \mathbf{E}(\widehat{\kappa}_j|\mathcal{B}_j^N), \quad d_i < j \leq I;$$

l'operazione di media viene effettuata quindi considerando a ogni passo solo le distribuzioni condizionate dato  $\mathcal{B}_{j-1}^R$  o dato  $\mathcal{B}_{j-1}^N$ . Nel ricampionamento per il DY  $j$  i dati effettivamente



osservati (i dati in  $\mathcal{D}_I$ ) fino al passo precedente sono mantenuti inalterati, e vengono generate “nuove osservazioni” solo relativamente al passo  $j$ , utilizzando la (4), la (26) e la (27); cioè:

$$\begin{aligned}\tilde{B}_{i,j} &= \alpha_j R_{i,j-1} + \beta_j \sqrt{R_{i,j-1}} \tilde{\varepsilon}_{i,j}^B, \\ \tilde{N}_{i,j} &= \varphi_j R_{i,j-1} + \psi_j \sqrt{R_{i,j-1}} \left( \tilde{\varepsilon}_{i,j}^N \sqrt{1 - \gamma_j^2 / (\beta_j^2 \psi_j^2)} + \tilde{\varepsilon}_{i,j}^B \gamma_j / (\beta_j \psi_j) \right), \\ \tilde{C}_{i,j} &= \kappa_j N_{i,j} + \rho_j \sqrt{N_{i,j}} \tilde{\varepsilon}_{i,j}^C,\end{aligned}$$

con le usuali ipotesi sulle  $\tilde{\varepsilon}_{i,j}^N, \tilde{\varepsilon}_{i,j}^B, \tilde{\varepsilon}_{i,j}^C$  (indipendenza, media nulla e varianza unitaria). I nuovi stimatori per  $\varphi_j, \alpha_j$  e  $\kappa_j$  utilizzano le osservazioni ricampionate nei numeratori, ma utilizzano i denominatori  $\sum_{i \in \mathcal{I}_j} R_{i,j-1}$ , per  $\varphi_j, \alpha_j$ , e i denominatori  $\sum_{i \in \mathcal{I}_j} N_{i,j}$ , per  $\kappa_j$ , come misure di volume fissate; hanno quindi la forma:

$$\begin{aligned}\tilde{\alpha}_j &= \frac{\sum_{i \in \mathcal{I}_j} \tilde{B}_{i,j}}{\sum_{i \in \mathcal{I}_j} R_{i,j-1}} = \alpha_j + \frac{\beta_j}{\sum_{i \in \mathcal{I}_j} R_{i,j-1}} \sum_{i \in \mathcal{I}_j} \sqrt{R_{i,j-1}} \tilde{\varepsilon}_{i,j}^B, \\ \tilde{\varphi}_j &= \frac{\sum_{i \in \mathcal{I}_j} \tilde{N}_{i,j}}{\sum_{i \in \mathcal{I}_j} R_{i,j-1}} = \varphi_j + \frac{\sqrt{\psi_j^2 - \gamma_j^2 / \beta_j^2}}{\sum_{i \in \mathcal{I}_j} R_{i,j-1}} \sum_{i \in \mathcal{I}_j} \sqrt{R_{i,j-1}} \tilde{\varepsilon}_{i,j}^N + \frac{\gamma_j / \beta_j}{\sum_{i \in \mathcal{I}_j} R_{i,j-1}} \sum_{i \in \mathcal{I}_j} \sqrt{R_{i,j-1}} \tilde{\varepsilon}_{i,j}^B, \\ \tilde{\kappa}_j &= \frac{\sum_{i \in \mathcal{I}_j} \tilde{C}_{i,j}}{\sum_{i \in \mathcal{I}_j} N_{i,j}} = \kappa_j + \frac{\rho_j}{\sum_{i \in \mathcal{I}_j} N_{i,j}} \sum_{i \in \mathcal{I}_j} \sqrt{N_{i,j}} \tilde{\varepsilon}_{i,j}^C.\end{aligned}\tag{50}$$

Tutti questi stimatori sono variabili aleatorie dato  $\mathcal{D}_I$ . Ovviamente lo stimatore per  $R_{ij}$  è dato da  $\tilde{f}_j = \tilde{\alpha}_j - \tilde{\varphi}_j$ .

Nel seguito le espressioni per gli stimatori del MSEP verranno ricavate per fasi successive sfruttando lo schema a catena, analogamente a quanto fatto per gli stimatori dei valori attesi. Dapprima si ricaveranno gli stimatori relativi al numero  $R_{i,j}$  dei riservati, si procederà quindi al calcolo degli stimatori per il numero  $N_{i,j}$  dei pagati chiusi e si ricaveranno infine gli stimatori per gli importi pagati  $C_{i,j}$ . Il vantaggio offerto da questa procedura è che le espressioni del MSEP per gli  $R_{i,j}$  sono essenzialmente già note, essendo state ricavate in [1] relativamente agli importi pagati cumulati; la seconda e la terza fase richiedono estensioni non complesse di queste formule.

Vale la pena di ricordare che lo schema a catena è adottato qui unicamente come metodologia di calcolo e non è assunto come modello per i numeri riservati. Precisamente, il modello per il processo R-da-R è quello espresso dalla (37) (ricavato combinando B-da-R e N-da-R) e non quello dato dalla (40). Quindi le definizioni:

$$f_j := \alpha_j - \varphi_j, \quad \sigma_j^2 := \beta_j^2 - 2\gamma_j + \psi_j^2,$$

e:

$$\hat{f}_j := \hat{\alpha}_j - \hat{\varphi}_j, \quad \hat{\sigma}_j^2 := \hat{\beta}_j^2 - 2\hat{\gamma}_j + \hat{\psi}_j^2,$$

sono adottate unicamente per comodità di notazione. In particolare, il fatto che lo stimatore  $\hat{\sigma}_j^2$  vada calcolato dagli stimatori  $\hat{\beta}_j^2, \hat{\psi}_j^2$  e  $\hat{\gamma}_j$  sta a ricordare che le stime dei parametri di varianza di  $R_{i,j}$  vanno ricavate dalle osservazioni di  $B_{i,j}$  e di  $N_{i,j}$  (tenendo conto della correlazione) e non dalle osservazioni di  $(B_{i,j} - N_{i,j})$ .

Tutte le espressioni “di cella” presentate nei paragrafi seguenti valgono in  $t = I$  per tutti gli anni  $(i, j)$  del triangolo futuro, cioè per  $i + j > I + 1$ . L’operatore  $\mathbf{\hat{E}}_{\mathcal{D}_I}$  indicherà, come si è detto, le stime di estimation error ricavate con l’approccio conditional resampling.

## 6.2 Numero dei riservati

Per la stima dell'estimation error conviene preliminarmente osservare che, dalla prima e dalla seconda delle (50), si ricavano facilmente le seguenti proprietà per gli stimatori  $\tilde{f}_j$ .

- 1R) Gli stimatori ricampionati  $\tilde{f}_2, \dots, \tilde{f}_I$  sono condizionatamente indipendenti rispetto a  $\mathcal{D}_I$ ;
- 2R)  $\mathbf{E}(\tilde{f}_j|\mathcal{D}_I) = f_j$ ,  $2 \leq j \leq I$ ;
- 3R)  $\mathbf{E}(\tilde{f}_j^2|\mathcal{D}_I) = f_j^2 + \sigma_j^2/S_{j-1}^{(R)}$ ,  $2 \leq j \leq I$ , avendo indicato con:

$$S_{j-1}^{(R)} := \sum_{i \in \mathcal{I}_j} R_{i,j-1},$$

le “somme di colonna” dei riservati osservati.

Per l'interpretazione delle formule di estimation error e di process error del numero di riservati sono utili, rispettivamente, le seguenti due osservazioni.

*Osservazione.* Per le (33), il termine  $\sigma_j^2/S_{j-1}^{(R)}$  nella 3R coincide con  $\mathbf{Var}(\hat{f}_j|\mathcal{B}_{j-1}^R)$ ; quindi  $\hat{\sigma}_j^2/S_{j-1}^{(R)} = (\hat{\psi}_j^2 - 2\hat{\gamma}_j + \hat{\beta}_j^2)/S_{j-1}^{(R)}$  fornisce la stima della varianza condizionata di  $\hat{f}_j = \hat{\alpha}_j - \hat{\varphi}_j$ , dato  $\mathcal{B}_{j-1}^R$ . ■

*Osservazione.* Se si considerano i fattori di sviluppo individuali dei riservati  $F_{i,j} = R_{i,j}/R_{i,j-1}$ , per la (37) si ha che la varianza condizionata di questi fattori ha la forma:

$$\mathbf{Var}(F_{i,j}|R_{i,j-1}) = \sigma_j^2/R_{i,j-1}.$$

Quindi  $\hat{\sigma}_j^2/\hat{R}_{i,j-1}$  fornisce una stima della varianza condizionata di  $F_{i,j}$ , dato  $R_{i,j-1}$ . ■

### 6.2.1 Stima delle process variance di cella

Lo stimatore della process variance  $\mathbf{Var}(R_{i,j}|\mathcal{D}_I)$  risulta avere la forma:

$$\widehat{\mathbf{Var}}(R_{i,j}|\mathcal{D}_I) = R_{i,d_i}^2 \sum_{k=d_i+1}^j \left( \prod_{h=d_i+1}^{k-1} \hat{f}_h^2 \right) \frac{\hat{\sigma}_k^2}{\hat{R}_{i,k-1}} \left( \prod_{h=k+1}^j \hat{f}_h^2 \right),$$

o, in termini di valori proiettati:

$$\widehat{\mathbf{Var}}(R_{i,j}|\mathcal{D}_I) = \hat{R}_{i,j}^2 \sum_{k=d_i+1}^j \frac{\hat{\sigma}_k^2/\hat{f}_k^2}{\hat{R}_{i,k-1}}, \quad (51)$$

dove  $\hat{R}_{i,j-1}$  per  $j > d_i + 1$  è dato dalla (45), e  $\hat{R}_{i,d_i} = R_{i,d_i}$ .

*Dimostrazione.* La linea dimostrativa è analoga a quella usata in [1], p. 529. ■

Dalla (51) si ricava anche la relazione ricorrente:

$$\widehat{\mathbf{Var}}(R_{i,j}|\mathcal{D}_I) = \widehat{\mathbf{Var}}(R_{i,j-1}|\mathcal{D}_I) \hat{f}_j^2 + \frac{\hat{\sigma}_j^2}{\hat{R}_{i,j-1}}.$$

### 6.2.2 Stima delle estimation variance di cella

Date le proprietà 1R-3R, lo stimatore dell'estimation error di  $R_{i,j}$  ha la forma:

$$\widehat{\mathbf{E}}_{\mathcal{D}_I} \left[ \left( \widehat{R}_{i,j} - \mathbf{E}[R_{i,j}|\mathcal{D}_I] \right)^2 \right] = R_{i,d_i}^2 \left[ \prod_{k=d_i+1}^j \left( \widehat{f}_k^2 + \frac{\widehat{\sigma}_k^2}{S_{k-1}^{(R)}} \right) - \prod_{k=d_i+1}^j \widehat{f}_k^2 \right], \quad (52)$$

o anche:

$$\widehat{\mathbf{E}}_{\mathcal{D}_I} \left[ \left( \widehat{R}_{i,j} - \mathbf{E}[R_{i,j}|\mathcal{D}_I] \right)^2 \right] = \widehat{R}_{i,j}^2 \left[ \prod_{k=d_i+1}^j \left( 1 + \frac{\widehat{\sigma}_k^2/\widehat{f}_k^2}{S_{k-1}^{(R)}} \right) - 1 \right].$$

La dimostrazione è analoga a quella sviluppata in [1], pp. 531-2.

### 6.3 Numero dei pagati chiusi

Dalla prima e dalla seconda delle (50) discendono le ulteriori proprietà per gli stimatori  $\widetilde{\varphi}_j$ , utili per gli stima dell'estimation error.

1N) Gli stimatori ricampionati  $\widetilde{\varphi}_2, \dots, \widetilde{\varphi}_I, \widetilde{f}_2, \dots, \widetilde{f}_I$  sono condizionatamente indipendenti rispetto a  $\mathcal{D}_I$ ;

2N)  $\mathbf{E}(\widetilde{\varphi}_j|\mathcal{D}_I) = \varphi_j$ ,  $2 \leq j \leq I$ ;

3N)  $\mathbf{E}(\widetilde{\varphi}_j^2|\mathcal{D}_I) = \varphi_j^2 + \psi_j^2/S_{j-1}^{(R)}$ ,  $2 \leq j \leq I$ ;

4N)  $\mathbf{E}(\widetilde{\varphi}_j \widetilde{f}_k|\mathcal{D}_I) = \varphi_j f_k$ ,  $2 \leq j, k \leq I$ ,  $j \neq k$ ;

5N)  $\mathbf{E}(\widetilde{\varphi}_j \widetilde{f}_j|\mathcal{D}_I) = \varphi_j f_j + (\gamma_j - \psi_j^2)/S_{j-1}^{(R)}$ ,  $2 \leq j \leq I$ .

Le seguenti osservazioni sono utili per l'interpretazione delle formule di estimation error e di process error del numero dei chiusi.

*Osservazione.* Per la seconda delle (33), il termine  $\psi_j^2/S_{j-1}^{(R)}$  nella 3N coincide con  $\mathbf{Var}(\widehat{\varphi}_j|\mathcal{B}_{j-1}^R)$ ; quindi  $\widehat{\psi}_j^2/S_{j-1}^{(R)}$  fornisce una stima per la varianza condizionata di  $\widehat{\varphi}_j$ , dato  $\mathcal{B}_{j-1}^R$ . Inoltre si può dimostrare che:

$$\mathbf{Cov}(\widehat{\alpha}_j, \widehat{\varphi}_j|\mathcal{B}_{j-1}^R) = \frac{\gamma_j}{S_{j-1}^R}.$$

Ne consegue che il termine  $(\gamma_j - \psi_j^2)/S_{j-1}^{(R)}$  nella 5N coincide con  $\mathbf{Cov}(\widehat{\varphi}_j, \widehat{f}_j|\mathcal{B}_{j-1}^R)$  e perciò  $(\widehat{\gamma}_j - \widehat{\psi}_j^2)/S_{j-1}^{(R)}$  può essere inteso come una stima della covarianza condizionata tra  $\widehat{\varphi}_j$  e  $\widehat{f}_j$ , dato  $\mathcal{B}_{j-1}^R$ . ■

*Osservazione.* Se si considerano i fattori di sviluppo individuali del processo dei pagati chiusi  $\Phi_{i,j} = N_{i,j}/R_{i,j-1}$ , si vede che per l'ipotesi di modello N3 la varianza condizionata di questi fattori ha la forma:

$$\mathbf{Var}(\Phi_{i,j}|R_{i,j-1}) = \psi_j^2/R_{i,j-1}.$$

Quindi  $\widehat{\psi}_j^2/\widehat{R}_{i,j-1}$  fornisce una stima della varianza condizionata di  $\Phi_{i,j}$ , dato  $R_{i,j-1}$ . Inoltre, come si controlla facilmente, per le ipotesi di modello N2 e N3 si ha:

$$\begin{aligned} \mathbf{Cov}(\Phi_{i,j}, F_{i,j}|R_{i,j-1}) &= \frac{1}{R_{i,j-1}^2} [\mathbf{Cov}(B_{i,j}, N_{i,j}|R_{i,j-1}) - \mathbf{Var}(N_{i,j}|R_{i,j-1})] \\ &= \frac{\widehat{\gamma}_j - \widehat{\psi}_j^2}{R_{i,j-1}}. \end{aligned}$$

Quindi  $(\widehat{\gamma}_j - \widehat{\psi}_j^2)/\widehat{R}_{i,j-1}$  si può intendere come la stima della covarianza condizionata, dato  $R_{i,j-1}$ , tra i fattori individuali  $\Phi_{i,j}$  e  $F_{i,j}$ . ■

### 6.3.1 Stima delle process variance di cella

Si ha:

$$\widehat{\mathbf{Var}}(N_{i,j}|\mathcal{D}_I) = \widehat{N}_{i,j}^2 \left( \frac{\widehat{\psi}_j^2/\widehat{\varphi}_j^2}{\widehat{R}_{i,j-1}} + \sum_{k=d_i+1}^{j-1} \frac{\widehat{\sigma}_k^2/\widehat{f}_k^2}{\widehat{R}_{i,k-1}} \right). \quad (53)$$

*Dimostrazione.* Infatti, dato che è  $N_j = \varphi_j R_{i,j-1} + \sqrt{R_{i,j-1}} \left( \widetilde{\varepsilon}_{i,j}^N \sqrt{\psi_j^2 - \gamma_j^2/\beta_j^2} + \widetilde{\varepsilon}_{i,j}^B \gamma_j/\beta_j \right)$ , si ha:

$$\begin{aligned} \mathbf{Var}(N_{i,j}|\mathcal{D}_I) &= \mathbf{E}(N_{i,j}^2|\mathcal{D}_I) - [\mathbf{E}(N_{i,j}|\mathcal{D}_I)]^2 \\ &= \varphi_j^2 \mathbf{E}(R_{i,j-1}^2|\mathcal{D}_I) + \psi_j^2 \mathbf{E}(R_{i,j-1}|\mathcal{D}_I) - \varphi_j^2 [\mathbf{E}(R_{i,j-1}|\mathcal{D}_I)]^2 \\ &= \varphi_j^2 \mathbf{Var}(R_{i,j-1}|\mathcal{D}_I) + \psi_j^2 \mathbf{E}(R_{i,j-1}|\mathcal{D}_I). \end{aligned} \quad (54)$$

Passando agli stimatori:

$$\widehat{\mathbf{Var}}(N_{i,j}|\mathcal{D}_I) = \widehat{\varphi}_j^2 \widehat{\mathbf{Var}}(R_{i,j-1}|\mathcal{D}_I) + \widehat{\psi}_j^2 \widehat{\mathbf{E}}(R_{i,j-1}|\mathcal{D}_I),$$

dove  $\widehat{\mathbf{Var}}(R_{i,j-1}|\mathcal{D}_I)$  è dato dalla (51) e  $\widehat{\mathbf{E}}(R_{i,j-1}|\mathcal{D}_I)$  è fornito dalla (44). Sostituendo, si ottiene la (53). ■

### 6.3.2 Stima delle process variance dei totali di AY

Dato che si tratta di tener conto delle covarianze “di riga”, si ha:

$$\widehat{\mathbf{Var}} \left( \sum_{j=d_i+1}^I N_{i,j} \middle| \mathcal{D}_I \right) = \sum_{j=d_i+1}^I \widehat{\mathbf{Var}}(N_{i,j}|\mathcal{D}_I) + 2 \sum_{d_i < j_1 < j_2 \leq I} \widehat{\mathbf{Cov}}(N_{i,j_1}, N_{i,j_2}|\mathcal{D}_I).$$

Risulta, per  $d_i < j_1 < j_2 \leq I$ :

$$\widehat{\mathbf{Cov}}(N_{i,j_1}, N_{i,j_2}|\mathcal{D}_I) = \widehat{N}_{i,j_1} \widehat{N}_{i,j_2} \left( \frac{(\widehat{\gamma}_{j_1} - \widehat{\psi}_{j_1}^2)/(\widehat{\varphi}_{j_1} \widehat{f}_{j_1})}{\widehat{R}_{i,j_1-1}} + \sum_{k=d_i+1}^{j_1-1} \frac{\widehat{\sigma}_k^2/\widehat{f}_k^2}{\widehat{R}_{i,k-1}} \right). \quad (55)$$

*Dimostrazione.* Sia infatti  $j_1 < j_2$ ; si ha:

$$\begin{aligned} \mathbf{Cov}(N_{i,j_1}, N_{i,j_2}|\mathcal{D}_I) &= \mathbf{E}[\mathbf{Cov}(N_{i,j_1}, N_{i,j_2}|R_{i,j_2-1}) | \mathcal{D}_I] \\ &\quad + \mathbf{Cov}[\mathbf{E}(N_{i,j_1}|R_{i,j_2-1}), \mathbf{E}(N_{i,j_2}|R_{i,j_2-1}) | \mathcal{D}_I] \\ &= \varphi_{j_2} \mathbf{Cov}(N_{i,j_1}, R_{i,j_2-1}|\mathcal{D}_I) \\ &= \varphi_{j_2} \{ \mathbf{E}[\mathbf{Cov}(N_{i,j_1}, R_{i,j_2-1}|R_{i,j_2-2}) | \mathcal{D}_I] \\ &\quad + \mathbf{Cov}[\mathbf{E}(N_{i,j_1}|R_{i,j_2-2}), \mathbf{E}(R_{i,j_2-1}|R_{i,j_2-2}) | \mathcal{D}_I] \} \\ &= \varphi_{j_2} f_{j_2-1} \mathbf{Cov}(N_{i,j_1}, R_{i,j_2-2}|\mathcal{D}_I) \\ &\quad \dots \\ &= \varphi_{j_2} f_{j_2-1} \cdots f_{j_1+1} \mathbf{Cov}(N_{i,j_1}, R_{i,j_1}|\mathcal{D}_I). \end{aligned}$$

Dato che  $R_{i,j} = B_{i,j} - N_{i,j}$  e ricordando la (30) e la (54), è anche:

$$\begin{aligned} \mathbf{Cov}(N_{i,j_1}, R_{i,j_1}|\mathcal{D}_I) &= \mathbf{Cov}(N_{i,j_1}, B_{i,j_1}|\mathcal{D}_I) - \mathbf{Var}(N_{i,j_1}|\mathcal{D}_I) \\ &= \gamma_{j_1} \mathbf{E}(R_{i,j_1-1}|\mathcal{D}_I) + \alpha_{j_1} \varphi_{j_1} \mathbf{Var}(R_{i,j_1-1}|\mathcal{D}_I) \\ &\quad - [\varphi_{j_1}^2 \mathbf{Var}(R_{i,j_1-1}|\mathcal{D}_I) + \psi_{j_1}^2 \mathbf{E}(R_{i,j_1-1}|\mathcal{D}_I)] \\ &= \varphi_{j_1} f_{j_1} \mathbf{Var}(R_{i,j_1-1}|\mathcal{D}_I) + (\gamma_{j_1} - \psi_{j_1}^2) \mathbf{E}(R_{i,j_1-1}|\mathcal{D}_I). \end{aligned}$$

Quindi:

$$\mathbf{Cov}(N_{i,j_1}, N_{i,j_2} | \mathcal{D}_I) = \varphi_{j_2} f_{j_2-1} \cdots f_{j_1+1} [\varphi_{j_1} f_{j_1} \mathbf{Var}(R_{i,j_1-1} | \mathcal{D}_I) + (\gamma_{j_1} - \psi_{j_1}^2) \mathbf{E}(R_{i,j_1-1} | \mathcal{D}_I)] .$$

Passando agli stimatori e sostituendo dalla (51) e dalla (45) si ottiene:

$$\widehat{\mathbf{Cov}}(N_{i,j_1}, N_{i,j_2} | \mathcal{D}_I) = \widehat{\varphi}_{j_2} \widehat{f}_{j_2-1} \cdots \widehat{f}_{j_1+1} \left( \widehat{\varphi}_{j_1} \widehat{f}_{j_1} \widehat{R}_{i,j_1-1}^2 \sum_{k=d_i+1}^{j_1-1} \frac{\widehat{\sigma}_k^2 / \widehat{f}_k^2}{\widehat{R}_{i,k-1}} + (\widehat{\gamma}_{j_1} - \widehat{\psi}_{j_1}^2) \widehat{R}_{i,j_1-1} \right)$$

Moltiplicando e dividendo il termine  $(\widehat{\gamma}_{j_1} - \widehat{\psi}_{j_1}^2) \widehat{R}_{i,j_1-1}$  per  $\widehat{\varphi}_{j_1} \widehat{f}_{j_1} \widehat{R}_{i,j_1-1}$  si ha la (55). ■

Date la (53) e la (55) la varianza di AY si può esprimere nella forma:

$$\widehat{\mathbf{Var}} \left( \sum_{j=d_i+1}^I N_{i,j} | \mathcal{D}_I \right) = \sum_{j_1, j_2=d_i+1}^I \widehat{N}_{i,j_1} \widehat{N}_{i,j_2} \left( \frac{\widehat{\zeta}_{j_1, j_2}^2}{\widehat{R}_{i, (j_1 \wedge j_2) - 1}} + \sum_{k=d_i+1}^{(j_1 \wedge j_2) - 1} \frac{\widehat{\sigma}_k^2 / \widehat{f}_k^2}{\widehat{R}_{i, k-1}} \right) ,$$

con  $(j \wedge k) := \min\{j, k\}$  e avendo definito:

$$\widehat{\zeta}_{j,k}^2 := \begin{cases} \widehat{\psi}_j^2 / \widehat{\varphi}_j^2 & \text{se } j = k, \\ (\widehat{\gamma}_{(j \wedge k)} - \widehat{\psi}_{(j \wedge k)}^2) / (\widehat{\varphi}_{(j \wedge k)} \widehat{f}_{(j \wedge k)}) & \text{se } j \neq k. \end{cases} \quad (56)$$

### 6.3.3 Stima della process variance dei totali globali

Per l'ipotesi di indipendenza tra AY non ci sono termini di covarianza tra righe per la process variance. Si ha quindi:

$$\widehat{\mathbf{Var}} \left( \sum_{i=2}^I \sum_{j=d_i+1}^I N_{i,j} | \mathcal{D}_I \right) = \sum_{i=2}^I \widehat{\mathbf{Var}} \left( \sum_{j=d_i+1}^I N_{i,j} | \mathcal{D}_I \right) .$$

### 6.3.4 Stima delle estimation variance di cella

Dato che è  $\widehat{N}_{i,j} = \widehat{\varphi}_j \widehat{R}_{i,j-1}$  e tenendo conto delle proprietà 1N-5N, l'espressione dello stimatore dell'estimation variance di  $N_{i,j}$  si ottiene con una ovvia estensione del procedimento che ha condotto alla (52). Risulta:

$$\widehat{\mathbf{E}}_{\mathcal{D}_I} \left[ \left( \widehat{N}_{i,j} - \mathbf{E}[N_{i,j} | \mathcal{D}_I] \right)^2 \right] = R_{i,d_i}^2 \left[ \left( \widehat{\varphi}_j^2 + \frac{\widehat{\psi}_j^2}{S_{j-1}^{(R)}} \right) \prod_{k=d_i+1}^{j-1} \left( \widehat{f}_k^2 + \frac{\widehat{\sigma}_k^2}{S_{k-1}^{(R)}} \right) - \widehat{\varphi}_j^2 \prod_{k=d_i+1}^{j-1} \widehat{f}_k^2 \right] ,$$

cioè:

$$\widehat{\mathbf{E}}_{\mathcal{D}_I} \left[ \left( \widehat{N}_{i,j} - \mathbf{E}[N_{i,j} | \mathcal{D}_I] \right)^2 \right] = \widehat{N}_{i,j}^2 \left[ \left( 1 + \frac{\widehat{\psi}_j^2 / \widehat{\varphi}_j^2}{S_{j-1}^{(R)}} \right) \prod_{k=d_i+1}^{j-1} \left( 1 + \frac{\widehat{\sigma}_k^2 / \widehat{f}_k^2}{S_{k-1}^{(R)}} \right) - 1 \right] .$$

### 6.3.5 Stima delle estimation variance dei totali di AY

Si ha:

$$\widehat{\mathbf{E}}_{\mathcal{D}_I} \left[ \left( \sum_{j=d_i+1}^I \left( \widehat{N}_{i,j} - \mathbf{E}[N_{i,j} | \mathcal{D}_I] \right) \right)^2 \right] = \sum_{j_1, j_2=d_i+1}^I \widehat{N}_{i,j_1} \widehat{N}_{i,j_2} \left[ \left( 1 + \frac{\widehat{\zeta}_{j_1, j_2}^2}{S_{j_1-1}^{(R)}} \right) \prod_{h=d_i+1}^{(j_1 \wedge j_2) - 1} \left( 1 + \frac{\widehat{\sigma}_h^2 / \widehat{f}_h^2}{S_{h-1}^{(R)}} \right) - 1 \right] .$$

Questa espressione si ottiene come caso particolare della analoga espressione (63) per gli importi ricavate nel successivo par. 6.4.5. Basta porre, per  $1 \leq j \leq I$ ,  $\kappa_j = 1$  e (con leggera estensione delle definizioni)  $\rho_j^2 = 0$ , il che equivale a supporre che ogni sinistro chiuso abbia un costo certo di 1 euro.

### 6.3.6 Stima dell'estimation variance dei totali globali

$$\widehat{\mathbf{E}}_{\mathcal{D}_I} \left[ \left( \sum_{i=2}^I \sum_{j=d_i+1}^I \left( \widehat{N}_{i,j} - \mathbf{E}[N_{i,j} | \mathcal{D}_I] \right) \right)^2 \right] =$$

$$\sum_{i,i_2=2}^I \sum_{j_1,j_2=d_{(i_1 \wedge i_2)}+1}^I \widehat{N}_{i_1,j_1} \widehat{N}_{i_2,j_2} \left[ \left( 1 + \frac{\widehat{\xi}_{j_1,j_2}^2}{S_{(j_1 \wedge j_2)}^{(R)} - 1} \right)^{(j_1 \wedge j_2)-1} \prod_{k=d_{i_1}+1}^{(j_1 \wedge j_2)-1} \left( 1 + \frac{\widehat{\sigma}_k^2 / \widehat{f}_k^2}{S_{k-1}^{(R)}} \right) - 1 \right].$$

Anche questa espressione si ottiene ponendo  $\kappa_j = 1$  e  $\rho_j^2 = 0$  per  $1 \leq j \leq I$  nell'espressione (67) che verrà ricavata nel par. 6.4.6.

## 6.4 Costi

Date le (50) si hanno le ulteriori proprietà per gli stimatori ricampionati.

1C) Gli stimatori ricampionati  $\widetilde{\kappa}_1, \dots, \widetilde{\kappa}_I, \widetilde{f}_2, \dots, \widetilde{f}_I, \widetilde{\varphi}_2, \dots, \widetilde{\varphi}_I$ , sono condizionatamente indipendenti rispetto a  $\mathcal{D}_I$ ;

2C)  $\mathbf{E}(\widetilde{\kappa}_j | \mathcal{D}_I) = \kappa_j$ ,  $1 \leq j \leq I$ ;

3C)  $\mathbf{E}(\widetilde{\kappa}_j^2 | \mathcal{D}_I) = \kappa_j^2 + \rho_j^2 / S_j^{(N)}$ ,  $1 \leq j \leq I$ , dove si sono indicate con:

$$S_j^{(N)} := \sum_{i \in \mathcal{I}_j} N_{i,j},$$

le somme di colonna dei pagati chiusi.

Le seguenti osservazioni sono utili a fini interpretativi.

*Osservazione.* Data la (7), il termine  $\rho_j^2 / S_j^{(N)}$  nella 3C coincide con  $\mathbf{Var}(\widehat{\kappa}_j | \mathcal{B}_j^N)$ ; quindi la quantità  $\widehat{\rho}_j^2 / S_j^{(N)}$  fornisce una stima della varianza condizionata di  $\widehat{\kappa}_j$ , dato  $\mathcal{B}_j^N$ . ■

*Osservazione.* Dall'ipotesi di modello K2 si ha che la varianza condizionata dei  $K_{i,j}$  ha la forma:

$$\mathbf{Var}(K_{i,j} | N_{i,j}) = \rho_j^2 / N_{i,j}.$$

Quindi  $\widehat{\rho}_j^2 / \widehat{N}_{i,j}$  fornisce una stima della varianza condizionata di  $K_{i,j}$ , dato  $N_{i,j}$ . ■

### 6.4.1 Stima delle process variance di cella

Si ha:

$$\widehat{\mathbf{Var}}(C_{i,j} | \mathcal{D}_I) = \widehat{C}_{i,j}^2 \left( \frac{\widehat{\rho}_j^2 / \widehat{\kappa}_j^2}{\widehat{N}_{i,j}} + \frac{\widehat{\psi}_j^2 / \widehat{\varphi}_j^2}{\widehat{R}_{i,j-1}} + \sum_{k=d_i+1}^{j-1} \frac{\widehat{\sigma}_k^2 / \widehat{f}_k^2}{\widehat{R}_{i,k-1}} \right). \quad (57)$$

*Dimostrazione.* Infatti si ha:

$$\begin{aligned} \mathbf{E}(C_{i,j}^2 | \mathcal{D}_I) &= \mathbf{E} \left[ \left( \kappa_j N_{i,j} + \rho_j \sqrt{N_{i,j}} \varepsilon_j^C \right)^2 \middle| \mathcal{D}_I \right] \\ &= \kappa_j^2 \mathbf{E}(N_{i,j}^2 | \mathcal{D}_I) + \rho_j^2 \mathbf{E}(N_{i,j} (\varepsilon_j^C)^2 | \mathcal{D}_I) + 2 \kappa_j \rho_j \mathbf{E}(N_{i,j}^{3/2} \varepsilon_j^C | \mathcal{D}_I) \\ &= \kappa_j^2 \mathbf{E}(N_{i,j}^2 | \mathcal{D}_I) + \rho_j^2 \mathbf{E}(N_{i,j} | \mathcal{D}_I). \end{aligned}$$

Dato che è  $\mathbf{E}(C_{i,j}|\mathcal{D}_I) = \kappa_j \mathbf{E}(N_{i,j}|\mathcal{D}_I)$ , si ottiene allora:

$$\begin{aligned}\mathbf{Var}(C_{i,j}|\mathcal{D}_I) &= \mathbf{E}(C_{i,j}^2|\mathcal{D}_I) - [\mathbf{E}(C_{i,j}|\mathcal{D}_I)]^2 \\ &= \kappa_j^2 \mathbf{E}(N_{i,j}^2|\mathcal{D}_I) + \rho_j^2 \mathbf{E}(N_{i,j}|\mathcal{D}_I) - \kappa_j^2 [\mathbf{E}(N_{i,j}|\mathcal{D}_I)]^2 \\ &= \kappa_j^2 \mathbf{Var}(N_{i,j}|\mathcal{D}_I) + \rho_j^2 \mathbf{E}(N_{i,j}|\mathcal{D}_I) .\end{aligned}$$

Passando agli stimatori:

$$\widehat{\mathbf{Var}}(C_{i,j}|\mathcal{D}_I) = \hat{\kappa}_j^2 \widehat{\mathbf{Var}}(N_{i,j}|\mathcal{D}_I) + \hat{\rho}_j^2 \hat{N}_{i,j} .$$

Moltiplicando e dividendo il secondo termine per  $\hat{\kappa}_j^2 \hat{N}_{i,j}$  e sostituendo dalla (53) e dalla (47), si ottiene la (57).  $\blacksquare$

#### 6.4.2 Stima della process variance delle OLL di AY

Per  $2 \leq i \leq I$ , si consideri il totale dei costi futuri dell'AY  $i$ , cioè le OLL  $L_i = \sum_{j=d_i+1}^I C_{i,j}$ ; la stima della process variance di  $L_i$  ha la forma:

$$\widehat{\mathbf{Var}}(L_i|\mathcal{D}_I) = \sum_{j=d_i+1}^I \widehat{\mathbf{Var}}(C_{i,j}|\mathcal{D}_I) + 2 \sum_{d_i < j_1 < j_2 \leq I} \widehat{\mathbf{Cov}}(C_{i,j_1}, C_{i,j_2}|\mathcal{D}_I) .$$

Per l'ipotesi di indipendenza tra le  $\varepsilon_{i,j}^C$  e le  $\varepsilon_{i,j}^N$ , per i termini di covarianza si ha  $\widehat{\mathbf{Cov}}(C_{i,j_1}, C_{i,j_2}|\mathcal{D}_I) = \hat{\kappa}_{j_1} \hat{\kappa}_{j_2} \widehat{\mathbf{Cov}}(N_{i,j_1}, N_{i,j_2}|\mathcal{D}_I)$ ; dalla (55) si ha quindi:

$$\widehat{\mathbf{Cov}}(C_{i,j_1}, C_{i,j_2}|\mathcal{D}_I) = \hat{C}_{i,j_1} \hat{C}_{i,j_2} \left( \frac{(\hat{\gamma}_{j_1} - \hat{\psi}_{j_1}^2)/(\hat{\varphi}_{j_1} \hat{f}_{j_1})}{\hat{R}_{i,j_1-1}} + \sum_{k=d_i+1}^{j_1-1} \frac{\hat{\sigma}_k^2/\hat{f}_k^2}{\hat{R}_{i,k-1}} \right) .$$

Si ottiene quindi:

$$\widehat{\mathbf{Var}}(L_i|\mathcal{D}_I) = \sum_{j_1, j_2=d_i+1}^I \hat{C}_{i,j_1} \hat{C}_{i,j_2} \left( \frac{\hat{\zeta}_{j_1, j_2}^2}{\hat{R}_{i, (j_1 \wedge j_2) - 1}} + \sum_{k=d_i+1}^{(j_1 \wedge j_2) - 1} \frac{\hat{\sigma}_k^2/\hat{f}_k^2}{\hat{R}_{i, k-1}} \right) , \quad (58)$$

dove  $\hat{\zeta}_{j,k}$  è definito nella (56).

#### 6.4.3 Stima della process variance delle OLL totali

Per l'ipotesi di indipendenza tra AY non ci sono termini di covarianza tra righe per la process variance. Per le OLL totali  $L = \sum_{i=2}^I L_i$  si ha quindi  $\widehat{\mathbf{Var}}(L|\mathcal{D}_I) = \sum_{i=2}^I \widehat{\mathbf{Var}}(L_i|\mathcal{D}_I)$ , cioè :

$$\widehat{\mathbf{Var}}(L|\mathcal{D}_I) = \sum_{i=2}^I \sum_{j_1, j_2=d_i+1}^I \hat{C}_{i,j_1} \hat{C}_{i,j_2} \left( \frac{\hat{\zeta}_{j_1, j_2}^2}{\hat{R}_{i, (j_1 \wedge j_2) - 1}} + \sum_{k=d_i+1}^{(j_1 \wedge j_2) - 1} \frac{\hat{\sigma}_k^2/\hat{f}_k^2}{\hat{R}_{i, k-1}} \right) . \quad (59)$$

#### 6.4.4 Stima delle estimation variance di cella

Si ha:

$$\hat{\mathbf{E}}_{\mathcal{D}_I} \left( \left[ \hat{C}_{i,j} - \mathbf{E}(C_{i,j}|\mathcal{D}_I) \right]^2 \right) = \hat{C}_{i,j}^2 \left[ \left( 1 + \frac{\hat{\rho}_j^2/\hat{\kappa}_j^2}{S_j^{(N)}} \right) \left( 1 + \frac{\hat{\psi}_j^2/\hat{\varphi}_j^2}{S_{j-1}^{(R)}} \right) \prod_{k=d_i+1}^{j-1} \left( 1 + \frac{\hat{\sigma}_k^2/\hat{f}_k^2}{S_{k-1}^{(R)}} \right) - 1 \right] . \quad (60)$$

*Dimostrazione.* Infatti, dato che è:

$$\widehat{C}_{i,j} = R_{i,d_i} \widehat{\kappa}_j \widehat{\varphi}_j \prod_{k=d_i+1}^{j-1} \widehat{f}_k, \quad \mathbf{E}(C_{i,j}|\mathcal{D}_I) = R_{i,d_i} \kappa_j \varphi_j \prod_{k=d_i+1}^{j-1} f_k,$$

si ha:

$$\begin{aligned} \left[ \widehat{C}_{i,j} - \mathbf{E}(C_{i,j}|\mathcal{D}_I) \right]^2 &= R_{i,d_i}^2 \left( \widehat{\kappa}_j^2 \widehat{\varphi}_j^2 \prod_{k=d_i+1}^{j-1} \widehat{f}_k^2 + \kappa_j^2 \varphi_j^2 \prod_{k=d_i+1}^{j-1} f_k^2 \right. \\ &\quad \left. - 2 (\widehat{\kappa}_j \kappa_j)(\widehat{\varphi}_j \varphi_j) \prod_{k=d_i+1}^{j-1} (\widehat{f}_k f_k) \right). \end{aligned}$$

Passando alla versione conditional resampling degli stimatori e prendendo l'aspettativa dato  $\mathcal{D}_I$ , si ottiene:

$$\begin{aligned} \widehat{\mathbf{E}}_{\mathcal{D}_I} \left( \left[ \widehat{C}_{i,j} - \mathbf{E}(C_{i,j}) \right]^2 \right) &= R_{i,d_i}^2 \left[ \left( \kappa_j^2 + \frac{\rho_j^2}{S_j^{(N)}} \right) \left( \varphi_j^2 + \frac{\psi_j^2}{S_{j-1}^{(R)}} \right) \prod_{k=d_i+1}^{j-1} \left( f_k^2 + \frac{\sigma_j^2}{S_{k-1}^{(R)}} \right) \right. \\ &\quad \left. - \kappa_j^2 \varphi_j^2 \prod_{k=d_i+1}^{j-1} f_k^2 \right]. \end{aligned}$$

Passando agli stimatori e moltiplicando e dividendo per il fattore  $(\widehat{\kappa}_j^2 \widehat{\varphi}_j^2 \prod_{k=d_i+1}^{j-1} \widehat{f}_k^2)$  si ottiene la (60). ■

#### 6.4.5 Stima dell'estimation variance delle OLL di AY

Si ha:

$$\widehat{\mathbf{E}}_{\mathcal{D}_I} \left( \left[ \widehat{L}_i - \mathbf{E}(L_i|\mathcal{D}_I) \right]^2 \right) = \sum_{j=d_i+1}^I \widehat{\mathbf{E}}_{\mathcal{D}_I} \left( \left[ \widehat{C}_{i,j} - \mathbf{E}(C_{i,j}|\mathcal{D}_I) \right]^2 \right) + 2 \sum_{d_i < j_1 < j_2 \leq I} \mathbf{EstCov}_{i;j_1,j_2}, \quad (61)$$

dove le covarianze tra celle hanno la forma:

$$\mathbf{EstCov}_{i;j_1,j_2} := \widehat{\mathbf{E}}_{\mathcal{D}_I} \left( \left[ \widehat{C}_{i,j_1} - \mathbf{E}(C_{i,j_1}|\mathcal{D}_I) \right] \left[ \widehat{C}_{i,j_2} - \mathbf{E}(C_{i,j_2}|\mathcal{D}_I) \right] \right).$$

Utilizzando le proprietà 1R-3R, 1N-5N e 1C-4C, risulta:

$$\mathbf{EstCov}_{i;j_1,j_2} = \widehat{C}_{i,j_1} \widehat{C}_{i,j_2} \left[ \left( 1 + \frac{(\widehat{\gamma}_{j_1} - \widehat{\psi}_{j_1}^2)/(\widehat{\varphi}_{j_1} \widehat{f}_{j_1})}{S_{j_1-1}^{(R)}} \right) \prod_{h=d_i+1}^{j_1-1} \left( 1 + \frac{\widehat{\sigma}_h^2/\widehat{f}_h^2}{S_{h-1}^{(R)}} \right) - 1 \right]. \quad (62)$$

*Dimostrazione.* Per la dimostrazione si usano ragionamenti analoghi a quelli che conducono alla (60). Si tratta di ricavare una stima del prodotto:

$$\left[ \widehat{C}_{i,j_1} - \mathbf{E}(C_{i,j_1}|\mathcal{D}_I) \right] \left[ \widehat{C}_{i,j_2} - \mathbf{E}(C_{i,j_2}|\mathcal{D}_I) \right], \quad d_i + 1 \leq j_1 < j_2 \leq I.$$



Si ha:

$$\begin{aligned}
& \left[ \widehat{C}_{i,j_1} - \mathbf{E}(C_{i,j_1} | \mathcal{D}_I) \right] \left[ \widehat{C}_{i,j_2} - \mathbf{E}(C_{i,j_2} | \mathcal{D}_I) \right] \\
&= R_{i,d_i}^2 \left( \widehat{\kappa}_{j_1} \widehat{\varphi}_{j_1} \prod_{h=d_i+1}^{j_1-1} \widehat{f}_h - \kappa_{j_1} \varphi_{j_1} \prod_{h=d_i+1}^{j_1-1} f_h \right) \left( \widehat{\kappa}_{j_2} \widehat{\varphi}_{j_2} \prod_{h=d_i+1}^{j_2-1} \widehat{f}_h - \kappa_{j_2} \varphi_{j_2} \prod_{h=d_i+1}^{j_2-1} f_h \right) \\
&= R_{i,d_i}^2 \left[ \left( \widehat{\kappa}_{j_1} \widehat{f}_{j_1} \widehat{\varphi}_{j_1} \prod_{h=d_i+1}^{j_1-1} \widehat{f}_h^2 \cdot \widehat{\kappa}_{j_2} \widehat{\varphi}_{j_2} \prod_{h=j_1+1}^{j_2-1} \widehat{f}_h \right) - \left( \kappa_{j_1} \widehat{f}_{j_1} \varphi_{j_1} \prod_{h=d_i+1}^{j_1-1} (f_h \widehat{f}_h) \cdot \widehat{\kappa}_{j_2} \widehat{\varphi}_{j_2} \prod_{h=j_1+1}^{j_2-1} \widehat{f}_h \right) \right. \\
&\quad \left. - \left( \widehat{\kappa}_{j_1} f_{j_1} \widehat{\varphi}_{j_1} \prod_{h=d_i+1}^{j_1-1} (f_h \widehat{f}_h) \cdot \kappa_{j_2} \varphi_{j_2} \prod_{h=j_1+1}^{j_2-1} f_h \right) + \left( \kappa_{j_1} f_{j_1} \varphi_{j_1} \prod_{h=d_i+1}^{j_1-1} f_h^2 \cdot \kappa_{j_2} \varphi_{j_2} \prod_{h=j_1+1}^{j_2-1} f_h \right) \right].
\end{aligned}$$

Passando agli stimatori a ricampionamento condizionato e prendendo l'aspettativa dato  $\mathcal{D}_I$ , il secondo e il quarto addendo si elidono. Tenendo poi conto anche delle 3R e 5N e passando agli stimatori, si ottiene:

$$\begin{aligned}
& \mathbf{E} \left( \left[ \widehat{C}_{i,j_1} - \mathbf{E}(C_{i,j_1} | \mathcal{D}_I) \right] \left[ \widehat{C}_{i,j_2} - \mathbf{E}(C_{i,j_2} | \mathcal{D}_I) \right] \middle| \mathcal{D}_I \right) \\
&= R_{i,d_i}^2 \left[ \left( \widehat{\varphi}_{j_1} \widehat{f}_{j_1} + \frac{(\widehat{\gamma}_{j_1} - \widehat{\psi}_{j_1}^2)}{S_{j_1-1}^{(R)}} \right) \prod_{h=d_i+1}^{j_1-1} \left( \widehat{f}_h^2 + \frac{\widehat{\sigma}_h^2}{S_{h-1}^{(R)}} \right) - \widehat{\varphi}_{j_1} \widehat{f}_{j_1} \prod_{h=d_i+1}^{j_1-1} \widehat{f}_h^2 \right] \times \\
&\quad \left( \widehat{\kappa}_{j_1} \widehat{\varphi}_{j_2} \widehat{\kappa}_{j_2} \prod_{h=j_1+1}^{j_2-1} \widehat{f}_h \right).
\end{aligned}$$

Moltiplicando e dividendo per  $(\widehat{\varphi}_{j_1} \widehat{f}_{j_1} \prod_{h=d_i+1}^{j_1-1} \widehat{f}_h^2) (\widehat{\kappa}_{j_1} \widehat{\varphi}_{j_2} \widehat{\kappa}_{j_2} \prod_{h=j_1+1}^{j_2-1} \widehat{f}_h)$ , si ha la (62). ■

Richiamando la (60), la (61) può essere espressa esplicitamente come:

$$\begin{aligned}
& \widehat{\mathbf{E}}_{\mathcal{D}_I} \left( \left[ \widehat{L}_i - \mathbf{E}(L_i | \mathcal{D}_I) \right]^2 \right) = \\
& \sum_{j_1, j_2 = d_i+1}^I \widehat{C}_{i,j_1} \widehat{C}_{i,j_2} \left[ \left( 1 + \frac{\widehat{\xi}_{j_1, j_2}^2}{S_{(j_1 \wedge j_2)}^{(N)}} \right) \left( 1 + \frac{\widehat{\xi}_{j_1, j_2}^2}{S_{(j_1 \wedge j_2)-1}^{(R)}} \right) \prod_{h=d_i+1}^{(j_1 \wedge j_2)-1} \left( 1 + \frac{\widehat{\sigma}_h^2 / \widehat{f}_h^2}{S_{h-1}^{(R)}} \right) - 1 \right], \quad (63)
\end{aligned}$$

dove  $\widehat{\xi}_{j,k}^2$  è definito nella (56) e  $\widehat{\xi}_{j,k}^2$  è definito come:

$$\widehat{\xi}_{j,k}^2 := \begin{cases} \widehat{\rho}_j^2 / \widehat{\kappa}_j^2 & \text{se } j = k, \\ 0 & \text{se } j \neq k. \end{cases}$$

#### 6.4.6 Stima dell'estimation variance delle OLL totali

Nonostante l'ipotesi di indipendenza tra AY, per l'estimation error delle OLL totali  $L := \sum_{i=2}^I L_i$  si hanno effetti di covarianza righe generati dai dati. Si ha:

$$\widehat{\mathbf{E}}_{\mathcal{D}_I} \left( \left[ \widehat{L} - \mathbf{E}(L | \mathcal{D}_I) \right]^2 \right) = \sum_{i=2}^I \widehat{\mathbf{E}}_{\mathcal{D}_I} \left( \left[ \widehat{L}_i - \mathbf{E}(L_i | \mathcal{D}_I) \right]^2 \right) + 2 \sum_{2 \leq i_1 < i_2 \leq I} \mathbf{EstCov}_{i_1, i_2}, \quad (64)$$

dove i termini di covarianza tra righe  $\mathbf{EstCov}_{i_1, i_2}$  hanno espressione:

$$\mathbf{EstCov}_{i_1, i_2} = \sum_{j_1=d_{i_1}+1}^I \sum_{j_2=d_{i_2}+1}^I \widehat{\mathbf{E}}_{\mathcal{D}_I} \left( \left[ \widehat{C}_{i_1, j_1} - \mathbf{E}(C_{i_1, j_1} | \mathcal{D}_I) \right] \left[ \widehat{C}_{i_2, j_2} - \mathbf{E}(C_{i_2, j_2} | \mathcal{D}_I) \right] \right).$$

Si trova, per  $2 \leq i_1 < i_2 \leq I$ :

$$\begin{aligned} \mathbf{EstCov}_{i_1, i_2} = & \sum_{j_1=d_{i_1}+1}^I \hat{C}_{i_1, j_1} \hat{C}_{i_2, j_1} \left[ \left( 1 + \frac{\hat{\rho}_{j_1}^2 / \hat{\kappa}_{j_1}^2}{S_{j_1}^{(N)}} \right) \left( 1 + \frac{\hat{\psi}_{j_1}^2 / \hat{\varphi}_{j_1}^2}{S_{j_1-1}^{(R)}} \right) \prod_{k=d_{i_1}+1}^{j_1-1} \left( 1 + \frac{\hat{\sigma}_k^2 / \hat{f}_k^2}{S_{k-1}^{(R)}} \right) - 1 \right] \\ & + 2 \sum_{d_{i_1} < j_1 < j_2 \leq J}^I \hat{C}_{i_1, j_1} \hat{C}_{i_2, j_2} \left[ \left( 1 + \frac{(\hat{\gamma}_{j_1} - \hat{\psi}_{j_1}^2) / (\hat{f}_{j_1} \hat{\varphi}_{j_1})}{S_{j_1-1}^{(R)}} \right) \prod_{k=d_{i_1}+1}^{j_1-1} \left( 1 + \frac{\hat{\sigma}_k^2 / \hat{f}_k^2}{S_{k-1}^{(R)}} \right) - 1 \right]. \end{aligned} \quad (65)$$

*Dimostrazione.* Anche qui si applicano ragionamenti analoghi a quelli usati per ottenere la (60) e la (62), tenendo conto che: i termini con  $j_2 \leq d_{i_1}$  danno contributo nullo; i termini con  $j_1 = j_2$  danno origine ai “termini quadrati”  $\hat{\varphi}_{j_1}^2$  e  $\hat{\kappa}_{j_1}^2$  (proprietà 3N e 3C); i termini con  $j_1 < j_2$  danno origine al “termine rettangolo”  $\hat{\varphi}_{j_1} \hat{f}_{j_1}$  (proprietà 5N); i termini con  $j_1 > j_2$  sono “simmetrici” rispetto a quelli con  $j_1 < j_2$ . ■

La (65) si può anche esprimere nella forma:

$$\mathbf{EstCov}_{i_1, i_2} = \sum_{j_1, j_2=d_{i_1}+1}^I \hat{C}_{i_1, j_1} \hat{C}_{i_2, j_2} \left[ \left( 1 + \frac{\hat{\xi}_{j_1, j_2}^2}{S_{(j_1 \wedge j_2)}^{(N)}} \right) \left( 1 + \frac{\hat{\zeta}_{j_1, j_2}^2}{S_{(j_1 \wedge j_2)-1}^{(R)}} \right) \prod_{k=d_{i_1}+1}^{(j_1 \wedge j_2)-1} \left( 1 + \frac{\hat{\sigma}_k^2 / \hat{f}_k^2}{S_{k-1}^{(R)}} \right) - 1 \right]. \quad (66)$$

Conglobando la (63) e la (65) nella (64) si ha:

$$\begin{aligned} \widehat{\mathbf{E}}_{\mathcal{D}_I} \left( \left[ \hat{L} - \mathbf{E}(L|\mathcal{D}_I) \right]^2 \right) = & \sum_{i_1, i_2=2}^I \sum_{j_1, j_2=d_{(i_1 \wedge i_2)}+1}^I \hat{C}_{i_1, j_1} \hat{C}_{i_2, j_2} \left[ \left( 1 + \frac{\hat{\xi}_{j_1, j_2}^2}{S_{(j_1 \wedge j_2)}^{(N)}} \right) \left( 1 + \frac{\hat{\zeta}_{j_1, j_2}^2}{S_{(j_1 \wedge j_2)-1}^{(R)}} \right) \prod_{k=d_{i_1}+1}^{(j_1 \wedge j_2)-1} \left( 1 + \frac{\hat{\sigma}_k^2 / \hat{f}_k^2}{S_{k-1}^{(R)}} \right) - 1 \right]. \end{aligned} \quad (67)$$

## 6.5 Espressioni globali per il MSEP

Dalle (59) e (67) la stima del MSEP delle OLL globali è data da:

$$\begin{aligned} \widehat{\text{mse}}_{\mathcal{D}_I}(\hat{L}) = & \widehat{\mathbf{Var}}(L|\mathcal{D}_I) + \widehat{\mathbf{E}}_{\mathcal{D}_I} \left( \left[ \hat{L} - \mathbf{E}(L|\mathcal{D}_I) \right]^2 \right) \\ = & \sum_{i=2}^I \sum_{j_1, j_2=d_i+1}^I \hat{C}_{i, j_1} \hat{C}_{i, j_2} \left( \frac{\hat{\xi}_{j_1, j_2}^2}{\hat{R}_{i, (j_1 \wedge j_2)-1}} + \sum_{k=d_i+1}^{(j_1 \wedge j_2)-1} \frac{\hat{\sigma}_k^2 / \hat{f}_k^2}{\hat{R}_{i, k-1}} \right) \\ & + \sum_{i_1, i_2=2}^I \sum_{j_1, j_2=d_{(i_1 \wedge i_2)}+1}^I \hat{C}_{i_1, j_1} \hat{C}_{i_2, j_2} \left[ \left( 1 + \frac{\hat{\xi}_{j_1, j_2}^2}{S_{(j_1 \wedge j_2)}^{(N)}} \right) \left( 1 + \frac{\hat{\zeta}_{j_1, j_2}^2}{S_{(j_1 \wedge j_2)-1}^{(R)}} \right) \prod_{k=d_{i_1}+1}^{(j_1 \wedge j_2)-1} \left( 1 + \frac{\hat{\sigma}_k^2 / \hat{f}_k^2}{S_{k-1}^{(R)}} \right) - 1 \right]. \end{aligned}$$

La stima del MSEP globale può essere espressa come somma delle stime di MSEP di AY più un termine di covarianza tra AY; dalle (58) e (63) si ha, per  $2 \leq i \leq I$ :

$$\begin{aligned} \widehat{\text{mse}}_{\mathcal{D}_I}(\widehat{L}_i) &= \widehat{\mathbf{Var}}(L_i|\mathcal{D}_I) + \widehat{\mathbf{E}}_{\mathcal{D}_I} \left( [\widehat{L}_i - \mathbf{E}(L_i|\mathcal{D}_I)]^2 \right) \\ &= \sum_{j_1, j_2=d_i+1}^I \widehat{C}_{i,j_1} \widehat{C}_{i,j_2} \left( \frac{\widehat{\xi}_{j_1,j_2}^2}{\widehat{R}_{i,(j_1 \wedge j_2)-1}} + \sum_{k=d_i+1}^{(j_1 \wedge j_2)-1} \frac{\widehat{\sigma}_k^2/\widehat{f}_k^2}{\widehat{R}_{i,k-1}} \right) \\ &\quad + \sum_{j_1, j_2=d_i+1}^I \widehat{C}_{i,j_1} \widehat{C}_{i,j_2} \left[ \left( 1 + \frac{\widehat{\xi}_{j_1,j_2}^2}{S_{(j_1 \wedge j_2)}^{(N)}} \right) \left( 1 + \frac{\widehat{\xi}_{j_1,j_2}^2}{S_{(j_1 \wedge j_2)-1}^{(R)}} \right) \prod_{h=d_i+1}^{(j_1 \wedge j_2)-1} \left( 1 + \frac{\widehat{\sigma}_h^2/\widehat{f}_h^2}{S_{h-1}^{(R)}} \right) - 1 \right]. \end{aligned}$$

Utilizzando la (66) si ha infine:

$$\begin{aligned} \widehat{\text{mse}}_{\mathcal{D}_I}(\widehat{L}) &= \sum_{i=2}^I \widehat{\text{mse}}_{\mathcal{D}_I}(\widehat{L}_i) \\ &\quad + 2 \sum_{2 \leq i_1 < i_2 \leq I} \sum_{j_1, j_2=d_{i_1}+1}^I \widehat{C}_{i_1,j_1} \widehat{C}_{i_2,j_2} \left[ \left( 1 + \frac{\widehat{\xi}_{j_1,j_2}^2}{S_{(j_1 \wedge j_2)}^{(N)}} \right) \left( 1 + \frac{\widehat{\xi}_{j_1,j_2}^2}{S_{(j_1 \wedge j_2)-1}^{(R)}} \right) \prod_{k=d_{i_1}+1}^{(j_1 \wedge j_2)-1} \left( 1 + \frac{\widehat{\sigma}_k^2/\widehat{f}_k^2}{S_{k-1}^{(R)}} \right) - 1 \right]. \end{aligned}$$

## 6.6 Espressioni globali di MSEP per la versione distribution free

Se si adotta la versione distribution free del modello descritta nel par. 5.2, si trova che tutte le espressioni per la componente di process variance del MSEP sono identiche a quelle per il modello time series ora ricavate. Per la componente di estimation error è però necessario riconsiderare la logica definitoria dell'errore di stima, dato che il punto di vista del conditional resampling non è applicabile nella modalità usata per il time series. Si può vedere che tutte le definizioni adottate in letteratura per questi tipi di modelli, a cominciare dal modello originario di Mack, fino al modello proposto recentemente in [3], conducono a una semplice modifica della struttura moltiplicativa tipica delle formule di estimation error del modello time series. In pratica le espressioni per l'approccio distribution free si ottengono dalle corrispondenti espressioni time series adottando l'approssimazione lineare  $\prod_{i=1}^n (1+x) \approx 1 + \sum_{i=1}^n x$ . Se si adotta questa approssimazione, anche le componenti di estimation error assumono una struttura additiva e si hanno le seguenti espressioni esplicite:

$$\begin{aligned} \widehat{\text{mse}}_{\mathcal{D}_I}(\widehat{L}) &= \widehat{\mathbf{Var}}(L|\mathcal{D}_I) + \widehat{\mathbf{E}}_{\mathcal{D}_I} \left( [\widehat{L} - \mathbf{E}(L|\mathcal{D}_I)]^2 \right) \\ &= \sum_{i=2}^I \sum_{j_1, j_2=d_i+1}^I \widehat{C}_{i,j_1} \widehat{C}_{i,j_2} \left( \frac{\widehat{\xi}_{j_1,j_2}^2}{\widehat{R}_{i,(j_1 \wedge j_2)-1}} + \sum_{k=d_i+1}^{(j_1 \wedge j_2)-1} \frac{\widehat{\sigma}_k^2/\widehat{f}_k^2}{\widehat{R}_{i,k-1}} \right) \\ &\quad + \sum_{i,i_2=2}^I \sum_{j_1, j_2=d_{(i_1 \wedge i_2)}+1}^I \widehat{C}_{i_1,j_1} \widehat{C}_{i_2,j_2} \left( \frac{\widehat{\xi}_{j_1,j_2}^2}{S_{(j_1 \wedge j_2)}^{(N)}} + \frac{\widehat{\xi}_{j_1,j_2}^2}{S_{(j_1 \wedge j_2)-1}^{(R)}} + \sum_{k=d_{i_1}+1}^{(j_1 \wedge j_2)-1} \frac{\widehat{\sigma}_k^2/\widehat{f}_k^2}{S_{k-1}^{(R)}} \right). \end{aligned}$$

Per  $2 \leq i \leq I$ :

$$\begin{aligned} \widehat{\text{mse}}_{\mathcal{D}_I}(\widehat{L}_i) &= \widehat{\mathbf{Var}}(L_i|\mathcal{D}_I) + \widehat{\mathbf{E}}_{\mathcal{D}_I} \left( [\widehat{L}_i - \mathbf{E}(L_i|\mathcal{D}_I)]^2 \right) \\ &= \sum_{j_1, j_2=d_i+1}^I \widehat{C}_{i,j_1} \widehat{C}_{i,j_2} \left( \frac{\widehat{\xi}_{j_1,j_2}^2}{\widehat{R}_{i,(j_1 \wedge j_2)-1}} + \sum_{k=d_i+1}^{(j_1 \wedge j_2)-1} \frac{\widehat{\sigma}_k^2/\widehat{f}_k^2}{\widehat{R}_{i,k-1}} \right) \\ &\quad + \sum_{j_1, j_2=d_i+1}^I \widehat{C}_{i,j_1} \widehat{C}_{i,j_2} \left( \frac{\widehat{\xi}_{j_1,j_2}^2}{S_{(j_1 \wedge j_2)}^{(N)}} + \frac{\widehat{\xi}_{j_1,j_2}^2}{S_{(j_1 \wedge j_2)-1}^{(R)}} + \sum_{h=d_i+1}^{(j_1 \wedge j_2)-1} \frac{\widehat{\sigma}_h^2/\widehat{f}_h^2}{S_{h-1}^{(R)}} \right). \end{aligned}$$

Eliminando i termini contenenti  $\hat{\xi}_{j,k}^2$  e sostituendo  $\hat{C}_{i,j}$  con  $\hat{N}_{i,j}$  (come si avrebbe ipotizzando un costo medio certo uguale a 1 per tutti i DY), questa espressione fornisce il MSEP di AY per i numeri di pagati chiusi. Con gli appropriati cambi di notazione (e di significato) si ha allora coincidenza con l'Estimator 4.1 in [3].

Infine si ha:

$$\widehat{\text{mse}}_{\mathcal{D}_I}(\hat{L}) = \sum_{i=2}^I \widehat{\text{mse}}_{\mathcal{D}_I}(\hat{L}_i) + 2 \sum_{2 \leq i_1 < i_2 \leq I} \sum_{j_1, j_2 = d_{i_1} + 1}^I \hat{C}_{i_1, j_1} \hat{C}_{i_2, j_2} \left( \frac{\hat{\xi}_{j_1, j_2}^2}{S_{(j_1 \wedge j_2)}^{(N)}} + \frac{\hat{\zeta}_{j_1, j_2}^2}{S_{(j_1 \wedge j_2) - 1}^{(R)}} + \sum_{k=d_{i_1} + 1}^{(j_1 \wedge j_2) - 1} \frac{\hat{\sigma}_k^2 / \hat{f}_k^2}{S_{k-1}^{(R)}} \right).$$

Anche qui, eliminando i termini con  $\hat{\xi}_{j,k}^2$  e sostituendo  $\hat{C}_{i,j}$  con  $\hat{N}_{i,j}$  si ottiene l'espressione per il MSEP globale per i numeri di chiusi; questa coincide formalmente con l'Estimator 4.2 in [3].

## 7 La stima del modello per simulazione con bootstrap

Si è già osservato che nel modello SFL le distribuzioni di probabilità di tutte le quantità rilevanti per le applicazioni di Solvency II, e i relativi momenti e percentili, possono essere ricavate solo con metodi numerici. L'approccio numerico standard per questo tipo di problemi è quello della simulazione Monte Carlo, comprensiva di una fase bootstrap per la generazione degli estimation error.

Per il modello SFL si tratta di applicare un bootstrap “tridimensionale”, dato che sono coinvolte le stime dei parametri di tre processi, quello dei costi medi e quello dei numeri, che è composto a sua volta da due processi concatenati ( $B$  e  $N$  nello schema in parallelo,  $N$  e  $R$  nello schema a catena). Per quanto riguarda il bootstrap sul processo dei numeri, inoltre, sono possibili diversi approcci alla definizione retrograda dei residui rispetto ai dati. Naturalmente, ciascuno di questi approcci corrisponde a uno dei possibili metodi di ricampionamento disponibili per ricavare l'estimation error, cui si è già accennato nel paragrafo 6.1; la definizione dei residui costituisce di fatto il cuore del resampling.

Dato che l'approccio qui adottato è quello del conditional resampling, lo schema di applicazione dei residui sarà uno schema condizionato. Inoltre, per migliorare l'efficienza numerica, si sceglierà l'approccio cosiddetto “parametrico”, in cui i residui sono ottenuti campionando variabili normali standard indipendenti. In sintesi, l'approccio usato è quello del “conditional parametric bootstrap”.

Nel seguito, dopo aver richiamato le ipotesi di base, sono elencate e descritte brevemente le fasi della procedura di simulazione del modello.

### 7.1 Ipotesi di base

Si è assunta l'uguaglianza  $J = I$  e si considerano gli importi pagati  $C_{i,j}$  espressi a valori correnti alla data  $t = I$ .

La simulazione può essere effettuata sia secondo lo schema in parallelo che quello a catena. Adottando lo schema in parallelo, si tratta di considerare il sistema ricorrente stocastico (43), che qui si richiama per comodità:

$$\begin{cases} B_{i,j} = \alpha_j R_{i,j-1} + \beta_j \sqrt{R_{i,j-1}} \varepsilon_j^B, \\ N_{i,j} = \varphi_j R_{i,j-1} + \psi_j \sqrt{R_{i,j-1}} \left( \varepsilon_{i,j}^N \sqrt{1 - \gamma_j^2 / (\beta_j^2 \psi_j^2)} + \varepsilon_{i,j}^B \gamma_j / (\beta_j \psi_j) \right), \\ R_{i,j} = B_{i,j} - N_{i,j}, \\ C_{i,j} = \kappa_j N_{i,j} + \rho_j \sqrt{N_{i,j}} \varepsilon_j^C. \end{cases} \quad (68)$$

Questo sistema è definito per  $2 \leq i \leq I$  e  $j > d_i$  a partire dal numero dei riservati  $R_{i,d_i}$  osservati alla data  $t = I$ . Simulando i termini di errore  $\varepsilon_{i,j}^B, \varepsilon_{i,j}^N$  e  $\varepsilon_{i,j}^C$  come v.a. indipendenti a media 0 e varianza 1, il sistema (68) produce valori simulati di tutte le terne  $\{C_{i,j}, N_{i,j}, R_{i,j}\} \in \mathcal{F}_I$ . Questa è la fase di proiezione con process error. In particolare siamo interessati ai valori  $C_{i,j}$  per  $i + j > I + 1$ .

Naturalmente i parametri che compaiono nelle (68) sono incogniti, e vanno quindi stimati dalle osservazioni  $\mathcal{D}_I$ . Come si è visto, gli stimatori appropriati per  $\kappa_j$  e  $\rho_j^2$  sono dati dalla (6) e (8), quelli per  $\alpha_j$  e  $\beta_j^2$  dalla (31) e (34), quelli per  $\varphi_j, \psi_j^2$  e  $\gamma_j$  dalla (32), (35) e (36), con i casi speciali  $\hat{\varphi}_I = \hat{\alpha}_I, \hat{\psi}_I = \hat{\beta}_I$  e  $\hat{\gamma}_I = \hat{\beta}_I^2$ . Per tener conto dell'incertezza legata alle stime dei parametri di media  $\hat{\kappa}_j, \hat{\varphi}_j$  e  $\hat{\alpha}_j$ , prima della fase di proiezione viene effettuata una procedura di ricampionamento condizionato dei dati. Questa è la fase di inclusione dell'estimation error con approccio bootstrap.

## 7.2 La procedura di simulazione

La generica iterazione della procedura di simulazione è strutturata come segue.

### 7.2.1 Generazione bootstrap delle pseudo stime

La componente di estimation error per i parametri  $\alpha_j, \varphi_j, \kappa_j$  viene ricavata generando pseudo stime con una procedura di bootstrap condizionato parametrico. Gli pseudo valori generati col bootstrap verranno indicati con “ $\sim$ ”.

#### Il bootstrap per i numeri

Per ricavare le pseudo stime dei parametri  $\alpha_j$  si generano dapprima gli pseudo fattori individuali  $\tilde{A}_{i,j}$  dei con-seguito dalla:

$$\tilde{A}_{i,j} = \hat{\alpha}_j + \frac{\hat{\beta}_j}{\sqrt{R_{i,j-1}}} \tilde{\epsilon}_{i,j}^B, \quad 2 \leq j \leq I, \quad i \in \mathcal{I}_j,$$

con  $\tilde{\epsilon}_{i,j}^B$  normali standard indipendenti. Si calcolano quindi le pseudo stime delle  $\alpha_j$  come:

$$\tilde{\alpha}_j = \sum_{i \in \mathcal{I}_j} \frac{R_{i,j-1}}{\sum_{k \in \mathcal{I}_j} R_{k,j-1}} \tilde{A}_{i,j}, \quad 2 \leq j \leq I.$$

Per le pseudo stime dei parametri  $\varphi_j$  si generano gli pseudo fattori individuali  $\tilde{\Phi}_{i,j}$  dei chiusi dalla:

$$\tilde{\Phi}_{i,j} = \hat{\varphi}_j + \frac{\sqrt{\hat{\psi}_j^2 - \hat{\gamma}_j^2 / \hat{\beta}_j^2}}{\sqrt{R_{i,j-1}}} \tilde{\epsilon}_{i,j}^N + \frac{\hat{\gamma}_j / \hat{\beta}_j}{\sqrt{R_{i,j-1}}} \tilde{\epsilon}_{i,j}^B, \quad 2 \leq j \leq I, \quad i \in \mathcal{I}_j,$$

con  $\tilde{\epsilon}_{i,j}^N$  normali standard indipendenti; si ricavano quindi le pseudo stime delle  $\varphi_j$  come:

$$\tilde{\varphi}_j = \sum_{i \in \mathcal{I}_j} \frac{R_{i,j-1}}{\sum_{k \in \mathcal{I}_j} R_{k,j-1}} \tilde{\Phi}_{i,j}, \quad 2 \leq j \leq I.$$

Come si vede, questi due passi riproducono numericamente le espressioni degli stimatori ricampionati  $\tilde{\alpha}_j$  e  $\tilde{\varphi}_j$  introdotti con la (50).

#### Il bootstrap per gli importi

Per ricavare le pseudo stime dei parametri  $\kappa_j$  si generano gli pseudo fattori individuali  $\tilde{K}_{i,j}$  dei costi medi dalla:

$$\tilde{K}_{i,j} = \hat{\kappa}_j + \frac{\hat{\rho}_j}{\sqrt{N_{i,j}}} \tilde{\epsilon}_{i,j}^C, \quad 1 \leq j \leq I, \quad i \in \mathcal{I}_j,$$

con  $\tilde{\epsilon}_{i,j}^C$  normali standard indipendenti. Si calcolano quindi le pseudo stime delle  $\kappa_j$  come:

$$\tilde{\kappa}_j = \sum_{i \in \mathcal{I}_j} \frac{N_{i,j}}{\sum_{k \in \mathcal{I}_j} N_{k,j}} \tilde{K}_{i,j}, \quad 1 \leq j \leq I.$$

In questo passo si produce quindi numericamente lo stimatore ricampionato  $\tilde{\kappa}_j$  nella (50).

### 7.2.2 Proiezione con process error dei numeri e degli importi futuri

#### Il triangolo dei numeri pagati futuri

Applicando i parametri  $\tilde{\varphi}_j$  e  $\tilde{\psi}_j$  all'ultima diagonale del triangolo dei numeri a riserva si ottiene il triangolo “futuro” degli pseudo numeri pagati, inclusivo del process error, tramite il seguente algoritmo.

Per  $2 \leq i \leq I$  e  $j = d_i + 1$ :

$$\begin{cases} \tilde{B}_{i,j} = \tilde{\alpha}_j R_{i,j-1} + \hat{\beta}_j \sqrt{R_{i,j-1}} \varepsilon_{i,j}^B, \\ \tilde{N}_{i,j} = \tilde{\varphi}_j R_{i,j-1} + \sqrt{R_{i,j-1}} \left( \varepsilon_{i,j}^N \sqrt{\hat{\psi}_j^2 - \hat{\gamma}_j^2 / \hat{\beta}_j^2} + \varepsilon_{i,j}^B \hat{\gamma}_j / \hat{\beta}_j \right); \end{cases}$$

per  $3 \leq i \leq I$  e  $d_i + 2 \leq j \leq I$ :

$$\begin{cases} \tilde{B}_{i,j} = \tilde{\alpha}_j \tilde{R}_{i,j-1} + \hat{\beta}_j \sqrt{\tilde{R}_{i,j-1}} \varepsilon_{i,j}^B, \\ \tilde{N}_{i,j} = \tilde{\varphi}_j \tilde{R}_{i,j-1} + \sqrt{\tilde{R}_{i,j-1}} \left( \varepsilon_{i,j}^N \sqrt{\hat{\psi}_j^2 - \hat{\gamma}_j^2 / \hat{\beta}_j^2} + \varepsilon_{i,j}^B \hat{\gamma}_j / \hat{\beta}_j \right), \\ \tilde{R}_{i,j} = \tilde{B}_{i,j} - \tilde{N}_{i,j}. \end{cases}$$

Naturalmente, come già osservato, è necessaria un'ipotesi specifica sulla distribuzione di probabilità dei termini di process error  $\varepsilon_{i,j}^B$  e  $\varepsilon_{i,j}^N$ . Scelte tipiche sono distribuzione uniforme standard o distribuzione normale standard.

#### Il triangolo degli importi pagati futuri

Applicando gli pseudo costi medi al triangolo futuro degli pseudo numeri  $\tilde{N}_{i,j}$  già costruito e aggiungendo il process error per i costi, si costruisce lo pseudo triangolo futuro degli importi:

$$\tilde{C}_{i,j} = \tilde{\kappa}_j \tilde{N}_{i,j} + \hat{\rho}_j \sqrt{\tilde{N}_{i,j}} \varepsilon_{i,j}^C, \quad i + j > I + 1,$$

dove le  $\varepsilon_{i,j}^C$  sono v.a. indipendenti, uniformi standard o normali standard.

### 7.2.3 La distribuzione empirica dei pagati futuri

I valori simulati degli importi pagati incrementali futuri  $\tilde{C}_{i,j}$  contengono quindi sia la componente di process error, generata dai termini di errore  $\varepsilon_{i,j}^N, \varepsilon_{i,j}^B$  e  $\varepsilon_{i,j}^C$ , sia la componente di estimation error relativa alle stime dei parametri  $\hat{\alpha}_j, \hat{\varphi}_j$  e  $\hat{\kappa}_j$ .

Raccogliendo i valori  $\tilde{C}_{i,j}$  generati in tutte le iterazioni si ottengono le distribuzioni empiriche dei pagati incrementali su tutte le celle future. In particolare le medie e le varianze campionarie – a livello di singola cella, per i totali per AY e per il totale sull'intero triangolo futuro – forniscono le stime empiriche dei valori attesi e dei MSEP corrispondenti, e sarà significativo confrontare queste stime campionarie con i valori forniti dalle espressioni analitiche ricavate nel paragrafo 6.

### 7.3 Includere gli effetti diagonali

Naturalmente i costi di rimborso  $C_{i,j}$  in tutte le celle future  $(i, j)$  si distinguono tra loro per la diversa data di esigibilità, o scadenza (*maturity*); precisamente, gli importi situati su una stessa diagonale, cioè i rimborsi  $\{C_{i,j}; i+j = I+1+\tau\}$ , saranno esigibili alla stessa scadenza  $t+\tau$ , con  $\tau = 1, \dots, I-1$ . Considerare quindi semplicemente la distribuzione della somma di  $C_{i,j}$ , per AY oppure globale, significa ignorare qualsiasi effetto “diagonale”, cioè qualsiasi effetto dovuto alla diversa scadenza. Si può dire che si adotta in questo caso un approccio di *loss reserving* di tipo “*undiscounted Liability-at-Maturity*” (“LM non scontato”), che è quello tradizionale. L’inclusione degli effetti diagonali è tuttavia opportuna, e anzi necessaria, nelle applicazioni di Solvency II. In particolare è necessario sia rivalutare i costi futuri con un modello per l’inflazione, sia considerarne il valore attuale; inoltre, dato che Solvency II considera un orizzonte temporale di un anno per il calcolo dell’SCR, è opportuno rappresentare gli impegni futuri, invece che con logica LM, come gli effettivi obblighi di fine anno, adottando l’approccio cosiddetto *Year-End Expectation* (YEE), o *Claims development Result* (CDR). L’approccio simulativo consente di rispondere a queste esigenze in modo semplice ed efficace.

#### 7.3.1 Aggiungere l’inflazione

Per passare a importi inflazionati è necessario introdurre un processo stocastico dell’inflazione di ramo. Il modo più semplice consiste nell’ipotizzare un effetto diagonale (inflazione dipendente solo dal tempo di calendario, identica per tutti i DY) assumendo che il fattore inflattivo  $p_\theta$  per la data  $\theta$  sia un moto browniano geometrico, descritto dall’equazione differenziale stocastica:

$$dp_\theta = \mu p_\theta dt + \omega p_\theta dZ_\theta ,$$

con  $\mu$  e  $\omega$  assegnati. In ogni iterazione, i pagati incrementali inflazionati  $\tilde{C}_{i,j}^*$  si ottengono da quelli non inflazionati  $\tilde{C}_{i,j}$  come:

$$\tilde{C}_{i,j}^* = \tilde{C}_{i,j} \exp \left[ (j - d_i)(\mu - \omega^2/2) + \omega \sum_{k=1}^{j-d_i} \varepsilon_k^p \right] , \quad i+j > I+1 ,$$

con  $\varepsilon_k^p$  normali standard indipendenti.

*Osservazione.* L’applicazione della volatilità all’inflazione futura non è un fatto scontato. Va infatti considerato che l’inflazione passata contiene di per sé una componente di volatilità (la volatilità “storica”), e che questa volatilità risulta quindi incorporata nei dati osservati, cioè nei costi storici. È verosimile supporre che il passaggio da costi storici a costi correnti, effettuato applicando la stima dell’inflazione storica, neutralizzi il trend inflattivo osservato ma mantenga nei dati, almeno in parte, la sua volatilità. Se questa volatilità implicita si scarica sulle stime dei coefficienti di varianza  $\rho_j^2$ , essa verrà poi inclusa nei costi futuri durante la fase di applicazione del process error, causando quindi una sopravvalutazione della variabilità inflattiva. ■

#### 7.3.2 Includere l’attualizzazione

Per esprimere i costi futuri in termini di valori attualizzati si tratta semplicemente di utilizzare la struttura per scadenza dei tassi di interesse risk-free  $i_\tau$  in vigore sul mercato in  $t = I$  per la scadenza  $\tau = 1, \dots, I-1$ . In ogni simulazione, dal triangolo futuro delle  $\tilde{C}_{i,j}^*$  inflazionate non scontate si passerà al triangolo dei corrispondenti valori scontati  $\tilde{M}_{i,j}$  attualizzando al tasso annuo  $\tau$  tutti gli importi sulla diagonale  $\tau$ ; cioè :

$$\tilde{M}_{i,j} = \tilde{C}_{i,j}^* (1 + i_{j-d_i})^{-(j-d_i)} , \quad i+j > I+1 .$$

Si passa così a un approccio LM (inflazionato) scontato.

### 7.3.3 Adottare l'approccio YEE

Gli obblighi di fine anno (year-end obligations) in  $t = I + 1$  relativi all'AY  $i = 2, \dots, I$ , sono dati da:

$$Z_i := C_{i,d_i+1}^* + V_i', \quad \text{con} \quad V_i' := \sum_{j=d_i+2}^J V_{i,j}',$$

dove  $V_{i,j}'$  è la stima in  $t = I + 1$  dei rimborsi residui  $C_{i,j}$  inflazionati e scontati. Gli impegni di fine anno sono cioè espressi come somma dell'importo liquidato e dell'importo posto a riserva (col metodo FL) per gli impegni residui. Naturalmente gli impegni di fine anno complessivi si otterranno sommando le  $Z_i$  su tutti gli AY aperti.

Per non introdurre un ulteriore processo stocastico non essenziale ai nostri scopi, conviene ipotizzare tassi di interesse deterministici, per cui la struttura dei tassi risk-free utilizzata per il discounting in  $t = I + 1$  è fornita semplicemente dalla struttura dei tassi forward a 1 anno  $i_{1,\tau}$  in  $t = I$ .

Per ricavare, in ogni simulazione, il valore degli impegni  $\tilde{Z}_i$ , oltre a calcolare la componente  $\tilde{C}_{i,d_i+1}^*$  di importi liquidati, si tratta di ricavare la stima FL  $\tilde{V}_{i,j}'$  degli importi riservati. A questo scopo si costruiscono i nuovi insiemi: di pseudo numeri riservati  $\tilde{\mathcal{R}}_{I+1}$ , di pseudo numeri pagati chiusi  $\tilde{\mathcal{N}}_{I+1}$  e di pseudo importi pagati  $\tilde{C}_{I+1}^*$ . Mentre  $\tilde{\mathcal{R}}_{I+1}$  e  $\tilde{\mathcal{N}}_{I+1}$  sono ottenuti aggiungendo ai dati effettivi dei numeri riservati  $\mathcal{R}_I$  e dei numeri chiusi  $\mathcal{N}_I$  le nuove diagonali di numeri riservati e chiusi simulati, l'insieme  $\tilde{C}_{I+1}^*$  è ricavato aggiungendo ai dati effettivi degli importi  $C_I$  la nuova diagonale di importi pagati simulati e rivalutando il tutto con l'inflazione simulata; riportando cioè tutti gli importi a costi correnti in  $t = I + 1$ . Da questi pseudo dati si ricavano poi le stime aggiornate  $\tilde{\varphi}_j', \tilde{\alpha}_j'$  e  $\tilde{\kappa}_j'$ ; le  $\tilde{V}_{i,j}'$  si ottengono quindi proiettando con l'FL, applicando il modello di inflazione per il periodo residuo e attualizzando coi tassi forward.

## 8 Un esempio numerico

### 8.1 I dati

Nelle tabelle 1-3 sono riportati i triangoli  $10 \times 10$  degli importi pagati, dei numeri di sinistri chiusi e dei numeri di sinistri riservati osservati in un tipico ramo danni negli anni dal 2000 al 2009. Per motivi di riservatezza tutti i dati sono stati modificati con l'applicazione di un fattore di scala. Per motivi di spazio gli importi pagati riportati sono arrotondati al migliaio di euro. Dato che il valore di  $R_{1,10}$  è nullo, si intende che il run off si ottiene in  $J = 10$  anni di sviluppo.

Nella tabella 4 è riportato il triangolo dei numeri di sinistri con seguito, ricavato dalla condizione di chiusura  $B_{i,j} = N_{i,j} + R_{i,j}$ .

### 8.2 I valori proiettati

Le stime dei parametri di media e deviazione standard del processo dei costi, del processo dei pagati chiusi dai riservati e dei con-seguito dai riservati sono riportate nelle tabelle 5, 6 e 7, rispettivamente. Data l'elevata instabilità delle stime dei parametri di covarianza, si è posto  $\hat{\gamma}_j = 0$  per  $2 \leq j < 10$ . Dato che è  $R_{1,10} = 0$ , non si hanno "effetti coda": tutti i con-seguito



$i$	$C_{i,1}$	$C_{i,2}$	$C_{i,3}$	$C_{i,4}$	$C_{i,5}$	$C_{i,6}$	$C_{i,7}$	$C_{i,8}$	$C_{i,9}$	$C_{i,10}$
1	81542	98809	42825	25434	18549	18799	12670	5000	4975	3090
2	85600	116317	47181	27269	20164	14231	10491	7244	6499	.
3	95970	121179	47600	26004	18259	7551	7784	8265	.	.
4	106067	113561	49785	24674	12197	7928	9197	.	.	.
5	106091	109411	45063	17300	10323	8129	.	.	.	.
6	108156	106263	38656	19027	9855	.	.	.	.	.
7	104880	108550	47423	22943	.	.	.	.	.	.
8	104497	99866	34288	.	.	.	.	.	.	.
9	104802	82812	.	.	.	.	.	.	.	.
10	101245	.	.	.	.	.	.	.	.	.

Tabella 1: Triangolo di pagati incrementali, espressi a costi correnti.

$i$	$N_{i,1}$	$N_{i,2}$	$N_{i,3}$	$N_{i,4}$	$N_{i,5}$	$N_{i,6}$	$N_{i,7}$	$N_{i,8}$	$N_{i,9}$	$N_{i,10}$
1	58121	30509	6924	3313	1744	994	556	348	132	69
2	53225	32868	7306	2982	1393	788	504	243	161	.
3	54156	30896	5671	2183	1083	566	326	219	.	.
4	55089	27136	4876	1785	882	459	289	.	.	.
5	53354	23422	3912	1645	821	469	.	.	.	.
6	51688	21485	4108	1849	991	.	.	.	.	.
7	47778	21195	5814	2173	.	.	.	.	.	.
8	48891	20702	3665	.	.	.	.	.	.	.
9	48789	17175	.	.	.	.	.	.	.	.
10	47257	.	.	.	.	.	.	.	.	.

Tabella 2: Triangolo dei numeri di pagati chiusi.

$i$	$R_{i,1}$	$R_{i,2}$	$R_{i,3}$	$R_{i,4}$	$R_{i,5}$	$R_{i,6}$	$R_{i,7}$	$R_{i,8}$	$R_{i,9}$	$R_{i,10}$
1	49722	13547	5739	2752	1489	770	401	161	59	0
2	50740	11762	4985	2503	1427	899	550	362	263	.
3	45310	8674	3841	1978	1182	753	502	361	.	.
4	38064	8516	3380	1825	1094	735	525	.	.	.
5	33550	7841	3519	1766	1099	696	.	.	.	.
6	32498	9022	3834	2271	1337	.	.	.	.	.
7	33779	11494	5317	3211	.	.	.	.	.	.
8	32991	7023	3939	.	.	.	.	.	.	.
9	25405	6370	.	.	.	.	.	.	.	.
10	24349	.	.	.	.	.	.	.	.	.

Tabella 3: Triangolo dei numeri dei riservati.

dell'ultimo DY vengono chiusi ( $N_{1,10} = B_{1,10}$ ); si ha quindi  $\hat{\varphi}_{10} = \hat{\alpha}_{10}$  e si è posto  $\hat{\psi}_{10} = \hat{\beta}_{10}$ . Nella tabella 8 sono anche riportate le stime della media della velocità di liquidazione  $\hat{v}_j$  e le stime dei parametri di media e deviazione standard del processo dei riservati dai riservati; ovviamente si ha  $\hat{v}_j = \hat{\varphi}_j / \hat{\alpha}_j$ ,  $\hat{f}_j = \hat{\alpha}_j - \hat{\varphi}_j$  e  $\hat{\sigma}_j^2 = \hat{\beta}_j^2 + \hat{\psi}_j^2$ .

Dai valori stimati dei parametri  $\hat{\varphi}_j$  e  $\hat{\alpha}_j$  sono state ricavate le stime puntuali dei numeri futuri  $N_{i,j}$  e  $R_{i,j}$ . I triangoli completati con questi valori proiettati (arrotondati a valori interi) sono riportati nelle tabelle 9 e 10. Nella tabella 11 è illustrato il triangolo degli importi pagati completato con i valori proiettati  $\hat{C}_{i,j}$  (sempre arrotondati alle migliaia di euro), ottenuti applicando ai valori proiettati  $\hat{N}_{i,j}$  le stime di costo medio  $\hat{\kappa}_j$ .

$i$	$B_{i,1}$	$B_{i,2}$	$B_{i,3}$	$B_{i,4}$	$B_{i,5}$	$B_{i,6}$	$B_{i,7}$	$B_{i,8}$	$B_{i,9}$	$B_{i,10}$
1	107843	44056	12663	6065	3233	1764	957	509	191	69
2	103965	44630	12291	5485	2820	1687	1054	605	424	..
3	99466	39570	9512	4161	2265	1319	828	580	.	..
4	93153	35652	8256	3610	1976	1194	814	.	.	..
5	86904	31263	7431	3411	1920	1165	.	.	.	..
6	84186	30507	7942	4120	2328	.	.	.	.	..
7	81557	32689	11131	5384	.	.	.	.	.	..
8	81882	27725	7604	.	.	.	.	.	.	..
9	74194	23545	.	.	.	.	.	.	.	..
10	71606	.	.	.	.	.	.	.	.	..

Tabella 4: Triangolo dei numeri dei con-seguito, ricavato dalla condizione di chiusura.

$j$	1	2	3	4	5	6	7	8	9	10
$\widehat{k}_j$	1.927	4.245	8.346	10.210	12.923	17.289	23.965	25.320	39.160	44.779
$\widehat{\rho}_j$	61.878	108.60	132.730	95.475	88.175	54.564	88.734	200.55	22.839	2.601

Tabella 5: Stima dei parametri di media e deviazione standard del processo dei costi.

$j$	2	3	4	5	6	7	8	9	10
$\widehat{\alpha}_j$	0.905	0.987	1.053	1.111	1.133	1.157	1.166	1.176	1.169
$\widehat{\beta}_j$	7.730	7.236	2.905	2.581	1.970	1.841	1.835	0.159	0.014

Tabella 6: Stima dei parametri di media e deviazione standard del processo dei  $B$ -da- $R$ .

$j$	2	3	4	5	6	7	8	9	10
$\widehat{\varphi}_j$	0.659	0.543	0.520	0.528	0.521	0.531	0.557	0.560	1.169
$\widehat{\psi}_j$	6.786	6.581	4.823	3.558	3.758	4.087	5.165	3.960	0.014

Tabella 7: Stima dei parametri di media e deviazione standard del processo dei  $N$ -da- $R$ .

$j$	2	3	4	5	6	7	8	9	10
$\widehat{v}_j$	0.728	0.550	0.494	0.475	0.460	0.459	0.478	0.476	1.000
$\widehat{f}_j$	0.246	0.444	0.533	0.583	0.612	0.627	0.608	0.616	0.000
$\widehat{\sigma}_j$	10.287	9.781	5.630	4.396	4.243	4.482	5.481	3.963	0.019

Tabella 8: Stima dei parametri dei processi “derivati”  $N$ -da- $B$  e  $R$ -da- $R$ .

Dai valori proiettati si ricavano le stime puntuali degli impegni in-essere (*outstanding loss liabilities*, OLL) per ciascun AY aperto e compressive. Le OLL per i numeri sono date dalla differenza tra i numeri ultimi e i numeri cumulati correnti; cioè sono ricavate come  $\widehat{L}_i^{(N)} := \sum_{j=d_i+1}^J \widehat{N}_{i,j}$ , per  $i = 2, \dots, I$ , e come  $\widehat{L}^{(N)} := \sum_{i=2}^I \widehat{L}_i^{(N)}$ .

Le OLL per gli importi dovrebbero essere ottenute dopo aver applicato alle stime dei costi futuri  $\widehat{C}_{i,j}$  un modello per l’inflazione futura di ramo. Inoltre, secondo il punto di vista Solvency II, questi valori inflazionati dovrebbero essere scontati con i tassi risk-free di mercato prima di essere sommati. Tuttavia, dato che interessa qui evidenziare solamente gli effetti caratteristici del “metodo Fisher-Lang”, nella presentazione dei risultati si prescinderà sia

$i$	$N_{i,1}$	$N_{i,2}$	$N_{i,3}$	$N_{i,4}$	$N_{i,5}$	$N_{i,6}$	$N_{i,7}$	$N_{i,8}$	$N_{i,9}$	$N_{i,10}$
1	58121	30509	6924	3313	1744	994	556	348	132	69
2	53225	32868	7306	2982	1393	788	504	243	161	308
3	54156	30896	5671	2183	1083	566	326	219	202	260
4	55089	27136	4876	1785	882	459	289	293	179	230
5	53354	23422	3912	1645	821	469	369	243	149	191
6	51688	21485	4108	1849	991	696	434	286	175	225
7	47778	21195	5814	2173	1695	974	608	400	245	314
8	48891	20702	3665	2050	1108	636	397	261	160	205
9	48789	17175	3458	1471	795	457	285	188	115	147
10	47257	16044	3256	1385	748	430	268	177	108	139

Tabella 9: Il triangolo dei numeri dei pagati chiusi completato con le stime puntuali  $\hat{N}_{i,j}$ .

$i$	$R_{i,1}$	$R_{i,2}$	$R_{i,3}$	$R_{i,4}$	$R_{i,5}$	$R_{i,6}$	$R_{i,7}$	$R_{i,8}$	$R_{i,9}$	$R_{i,10}$
1	49722	13547	5739	2752	1489	770	401	161	59	0
2	50740	11762	4985	2503	1427	899	550	362	263	0
3	45310	8674	3841	1978	1182	753	502	361	222	0
4	38064	8516	3380	1825	1094	735	525	319	197	0
5	33550	7841	3519	1766	1099	696	436	265	163	0
6	32498	9022	3834	2271	1337	819	513	312	192	0
7	33779	11494	5317	3211	1870	1146	718	437	269	0
8	32991	7023	3939	2098	1222	748	469	285	176	0
9	25405	6370	2826	1505	877	537	336	205	126	0
10	24349	5997	2661	1417	826	506	317	193	119	0

Tabella 10: Il triangolo dei numeri dei riservati completato con le stime puntuali  $\hat{R}_{i,j}$ .

$i$	$C_{i,1}$	$C_{i,2}$	$C_{i,3}$	$C_{i,4}$	$C_{i,5}$	$C_{i,6}$	$C_{i,7}$	$C_{i,8}$	$C_{i,9}$	$C_{i,10}$
1	81542	98809	42825	25434	18549	18799	12670	5000	4975	3090
2	85600	116317	47181	27269	20164	14231	10491	7244	6499	13773
3	95970	121179	47600	26004	18259	7551	7784	8265	7920	11640
4	106067	113561	49785	24674	12197	7928	9197	7410	7007	10299
5	106091	109411	45063	17300	10323	8129	8850	6155	5820	8554
6	108156	106263	38656	19027	9855	12037	10412	7242	6848	10064
7	104880	108550	47423	22943	21909	16840	14566	10131	9580	14080
8	104497	99866	34288	20927	14314	11003	9517	6619	6259	9199
9	104802	82812	28858	15016	10271	7895	6829	4750	4491	6601
10	101245	68106	27169	14137	9670	7433	6429	4472	4228	6214

Tabella 11: Il triangolo degli importi pagati completato con le stime puntuali  $\hat{C}_{i,j}$ .

dall'inflazione futura che dal discounting. Con questa scelta anche le OLL per gli importi sono date dalla differenza tra i costi ultimi e i pagati cumulati correnti; quindi sono calcolate come  $\hat{L}_i := \sum_{j=d_i+1}^J \hat{C}_{i,j}$ , per  $i = 2, \dots, I$ , e come  $\hat{L} := \sum_{i=2}^I \hat{L}_i$ .

I valori delle stime di OLL  $\hat{L}^{(N)}$  e  $\hat{L}$  ottenuti applicando le formule dell'FL sono riportati nella tabella 12. Nella tabella sono anche riportati i corrispondenti valori ottenuti dalla procedura di simulazione, lanciata con 10 000 iterazioni; sono anche evidenziati gli scarti percentuali tra

valori simulati e valori teorici.

$i$	numeri: $\hat{L}^{(N)}$			importi: $\hat{L}$		
	simulati	teorici	(%)	simulati	teorici	(%)
2	308	308	0.00	13774	13773	0.01
3	462	462	0.03	19561	19559	0.01
4	703	702	0.13	24777	24716	0.25
5	953	952	0.09	29503	29379	0.42
6	1817	1816	0.02	46601	46603	-0.00
7	4234	4236	-0.06	86931	87106	-0.20
8	4818	4818	0.00	77915	77839	0.10
9	6922	6915	0.11	84876	84709	0.20
10	22561	22554	0.03	148148	147857	0.20
globale	42777	42762	0.03	532085	531543	0.10

Tabella 12: Valori attesi delle OLL, simulati e teorici.

È interessante confrontare le proiezioni ottenute con l'FL con quelle fornite dal chain ladder. Il confronto può essere effettuato in almeno due modi diversi. Si può pensare a un modello frequency-severity che stima i costi medi nello stesso modo dell'FL, ma che stima i numeri ultimi applicando l'algoritmo chain ladder al triangolo dei numeri chiusi  $N_{i,j}$  cumulati. In questo caso il confronto tra i due approcci si può fare comparando semplicemente le stime dei numeri. Da altro punto di vista, si può pensare di applicare il chain ladder direttamente al triangolo delle paid losses  $C_{i,j}$  cumulate; sarà allora espressivo il confronto tra i costi ultimi. Questi confronti tra stime di OLL di numeri e di importi prodotti dai due metodi (con formula chiusa) sono riportati nella tabella 13. Risulta che, su questi dati, il metodo FL produce una stima di OLL globali più bassa sia di quella ottenuta da un modello frequency-severity con chain ladder sui numeri, sia di quella ricavata con una classica applicazione del chain ladder direttamente agli importi.

$i$	numeri			importi		
	CL	FL	(%)	CL	FL	(%)
2	67	308	359.97	3354	13773	310.64
3	202	462	128.70	9428	19559	107.46
4	440	702	59.28	16391	24716	50.79
5	773	952	23.23	25199	29379	16.59
6	1310	1816	38.70	35788	46603	30.22
7	2262	4236	87.27	53237	87106	63.62
8	4157	4818	15.89	70653	77839	10.17
9	8417	6915	-17.85	107015	84709	-20.84
10	31525	22554	-28.46	227226	147857	-34.93
globale	49152	42762	-13.00	548291	531543	-3.05

Tabella 13: Confronti con valori proiettati chain-ladder (valori teorici).

### 8.3 I prediction error

Nelle applicazioni pratiche del modello SFL l'approccio simulativo risulta essenziale. In questo esempio si è scelto di generare i termini di errore  $\varepsilon_{i,j}^B, \varepsilon_{i,j}^N$  e  $\varepsilon_{i,j}^C$  come campionamenti indipendenti di v.a. uniformi standardizzate. Per verificare la correttezza della procedura di simulazione con bootstrap è stato effettuato un confronto tra le stime di prediction error

$\widehat{\text{mse}}(\widehat{C}_{i,j})$  dei costi calcolati con le formule chiuse ricavate nel paragrafo 6 e i corrispondenti valori prodotti per simulazione. Nella tabella 14 sono riportati gli scarti percentuali  $\Delta_{i,j}$  tra valori simulati e valori teorici di  $[\widehat{\text{mse}}(\widehat{C}_{i,j})]^{1/2}$  in ogni cella futura  $(i,j)$ . Nella tabella 15 sono inoltre confrontati i valori di deviazioni standard (radice del MSE) dei totali di AY e globali degli importi pagati, ottenuti per simulazione e con formula chiusa.

$i$	$\Delta_{i,2}$	$\Delta_{i,3}$	$\Delta_{i,4}$	$\Delta_{i,5}$	$\Delta_{i,6}$	$\Delta_{i,7}$	$\Delta_{i,8}$	$\Delta_{i,9}$	$\Delta_{i,10}$
2	.	.	.	.	.	.	.	.	-0.46
3	.	.	.	.	.	.	.	-0.04	-0.36
4	.	.	.	.	.	.	1.18	1.10	0.10
5	.	.	.	.	.	1.00	1.54	2.48	1.50
6	.	.	.	.	-0.77	0.81	2.30	3.92	2.16
7	.	.	.	0.31	0.14	0.64	1.43	2.20	1.09
8	.	.	-0.30	0.04	-0.19	1.77	3.45	2.27	1.38
9	.	-0.22	-0.60	0.21	0.75	1.75	2.03	3.21	3.73
10	0.59	-0.01	0.37	-0.43	1.41	0.85	4.85	3.52	0.61

Tabella 14: Scarti percentuali simulato-teorico di prediction error dei costi  $C_{i,j}$  futuri (deviazioni standard).

$i$	importi: $\widehat{\text{mse}}^{1/2}(\widehat{L})$		
	simulati	teorici	(%)
2	109	109	-0.46
3	1376	1381	-0.35
4	6411	6392	0.30
5	7129	7072	0.81
6	9552	9444	1.14
7	14389	14472	-0.58
8	14278	14245	0.23
9	19777	19618	0.81
10	30395	30616	-0.72
globale	51050	51345	-0.57

Tabella 15: Deviazioni standard delle OLL, simulate e teoriche.

Nella tabella 16 sono elencate le deviazione standard, e i relativi coefficienti di variazione, delle OLL-numeri; per ogni generazione e per il globale, è riportata la scomposizione del prediction error in estimation error e process error. Dato che si considera la variabilità della somma dei numeri relativi a anni di calendario diversi, si tratta di un approccio di tipo Liability-at-Maturity (LM). Le analoghe stime di variabilità per gli importi sono riportate nella tabella 17. Si tratta sempre di OLL ottenute come somma di importi futuri (non scontati e non inflazionati) esigibili a diverse date future; si sta quindi considerando un approccio LM non scontato (e non inflazionato).

Come nel caso deterministico, è interessante confrontare i risultati dell'SFL con quelli forniti da una delle possibili versioni stocastiche del chain ladder. Qui si è considerata l'estensione BMW del modello DFCL di Mack, che ha la struttura di ipotesi più vicina a quella del'SFL. Nella tabella 18 i coefficienti di variazione della tabella 17 sono confrontati con i corrispondenti coefficienti di variazione ricavati applicando il BMW al triangolo degli importi cumulati. Come si è visto (tabella 12), su questi dati il valore della media  $\widehat{L}^{(N)}$  fornito dal BMW è più elevato di quello prodotto dall'SFL; tuttavia, risulta che il BMW fornisce complessivamente un livello più basso della deviazione standard unitaria globale.

SFL - Numero sinistri chiusi - Approccio LM						
<i>i</i>	prediction		estimation		process	
	Std	Cv(%)	Std	Cv(%)	Std	Cv(%)
2	1	0.17	0	0.15	0	0.07
3	17	3.69	11	2.38	13	2.86
4	74	10.59	39	5.59	64	9.09
5	106	11.10	48	5.01	95	9.96
6	168	9.26	76	4.20	151	8.32
7	311	7.34	150	3.54	274	6.47
8	369	7.65	136	2.82	340	7.06
9	808	11.68	227	3.29	774	11.17
10	1567	6.95	407	1.80	1511	6.70
globale	1919	4.49	748	1.75	1761	4.12

Tabella 16: Deviazioni standard e coefficienti di variazione dei numeri (valori simulati).

SFL - Importi pagati - Approccio LM						
<i>i</i>	prediction		estimation		process	
	Std	Cv(%)	Std	Cv(%)	Std	Cv(%)
2	109	0.79	98	0.71	47	0.34
3	1376	7.03	886	4.53	1061	5.42
4	6411	25.88	3315	13.40	5440	21.99
5	7129	24.16	3305	11.24	6279	21.31
6	9552	20.50	4504	9.65	8281	17.79
7	14389	16.55	7394	8.48	12268	14.12
8	14278	18.33	5666	7.27	12998	16.69
9	19777	23.30	5951	7.02	18735	22.08
10	30395	20.52	8367	5.66	29173	19.70
globale	51050	9.59	30699	5.77	40921	7.69

Tabella 17: Deviazioni standard e coefficienti di variazione degli importi (valori simulati).

SFL <i>vs</i> BMW - Importi pagati (LM, Cv(%))						
<i>i</i>	prediction		estimation		process	
	SFL	BMW	SFL	BMW	SFL	BMW
2	0.79	18.40	0.71	13.28	0.34	12.74
3	7.03	12.24	4.53	7.71	5.42	9.51
4	25.88	12.39	13.40	6.79	21.99	10.37
5	24.16	13.98	11.24	6.61	21.31	12.32
6	20.50	20.84	9.65	8.70	17.79	18.93
7	16.55	18.46	8.48	7.59	14.12	16.83
8	18.33	15.21	7.27	5.84	16.69	14.04
9	23.30	11.62	7.02	4.19	22.08	10.84
10	20.52	13.74	5.66	4.53	19.70	12.97
globale	9.59	7.62	5.77	4.10	7.69	6.43

Tabella 18: Confronto con chain ladder stocastico (BMW).

Nella tabella 19 sono illustrati i valori di deviazione standard degli impegni (importi) di fine anno, cioè delle OLL stimate con l'approccio YEE. Si conferma il risultato tipico, secondo il quale il metodo YEE tende a fornire una variabilità stimata più bassa rispetto al metodo

LM.

SFL - Importi pagati - Approccio YEE					
prediction		estimation		process	
Std	Cv(%)	Std	Cv(%)	Std	Cv(%)
41458	7.79	20152	3.79	36154	6.79

Tabella 19: Deviazioni standard e coefficienti di variazione degli importi secondo l'approccio YEE (valori simulati).

Dato che il modello SFL fornisce per simulazione l'intera distribuzione empirica di probabilità delle OLL (comprensiva sia del process risk che dell'estimation risk), è interessante ricavare il percentile degli impegni in-essere a un fissato livello  $p$  di probabilità, e calcolare quindi il corrispondente "risk capital di riserva" come *unexpected loss*, sottraendo cioè il valore degli impegni attesi. Per  $p = 99.5\%$  si ottiene così una misura di risk capital che può essere interpretata come il *reserve SCR* fornito da un modello interno nel quadro Solvency II. Nella tabella 20 questi valori di SCR, espressi per unità di OLL attese, sono forniti sia con approccio LM che YEE, e confrontati con i corrispondenti coefficienti di variazione.

SFL - Importi pagati			
LM		YEE	
Cv(%)	SCR(%)	Cv(%)	SCR(%)
9.59	27.36	7.79	20.68

Tabella 20: Confronto tra CV e SCR unitario, secondo l'approccio LM e l'approccio YEE.

Tutti i risultati precedenti sono stati ottenuti utilizzando per le stime dei parametri tutti i dati a disposizione, cioè ponendo  $\mathcal{I}_j = \{1, \dots, d_j\}$  per tutti i  $j$ . Il modello può essere applicato ponendo invece  $\mathcal{I}_j = \{\max\{1, d_j - 2\}, \dots, d_j\}$ , effettuando cioè le stime solo sui tre AY più recenti. Nella tabella 21 sono riportati i valori di media e deviazione standard delle OLL-importi, con la scomposizione in estimation e process error, ottenuti con numero di osservazioni ridotto. Dal confronto con i risultati prodotti dal campione completo si evidenzia un aumento delle OLL attese (i valori riportati sono quelli teorici), ma anche del livello di rischiosità, sia secondo l'LM che l'YEE. L'aumento di rischiosità secondo l'LM è essenzialmente dovuto all'aumento dell'estimation error (si ha anzi una leggera diminuzione del process error). L'YEE produce un aumento di rischiosità più marcato, dovuto sia a un incremento di estimation error che di process error.

	SFL - Importi pagati - Approccio YEE						
	$\hat{L}$	prediction		estimation		process	
		Std	Cv(%)	Std	Cv(%)	Std	Cv(%)
LM - ultimi 3 AY:	551285	53567	9.71	36142	6.55	39677	7.20
LM - tutti gli AY:	531543	51050	9.59	30699	5.77	40921	7.69
YEE - ultimi 3 AY:	551285	49156	8.77	26715	4.77	41296	7.37
YEE - tutti gli AY:	531543	41458	7.79	20152	3.79	36154	6.79

Tabella 21: Effetti della riduzione del numero di osservazioni usate per la stima.

## Riferimenti bibliografici

- [1] M. Buchwalder, M., Bühlmann, H., Merz, M, Wüthrich, M., *The Mean Square Error of Prediction in the Chain Ladder Reserving Method (Mack and Murphy Revisited)*, ASTIN Bulletin, **36**, 521-542, 2006.
- [2] Bühlmann, H., Gisler A., *A Course on Credibility Theory and its Applications*, Springer, Berlin 2005.
- [3] Dahms, R., *A Loss Reserving Method for Incomplete Data*, Mitteilungen-Bulletin SAV, 1-2, pp.127-148, 2008.
- [4] De Felice, M., Moriconi, *Risk based capital in P&C Loss reserving or stressing the triangle*, Research Group on “Insurance Companies and Pension Funds”, Working Paper n. 1, Rome, December 2003.
- [5] Fisher, W.H., Lange, J. T., *Loss Reserve Testing: a report year approach*, Proceedings of the Casualty Actuarial Society, Vol. LX, pp.189-207, 1973.
- [6] Mack, T. *Distribution-free Calculation of the Standard Error of Chain Ladder Reserve Estimates*, ASTIN Bulletin, **23**, 213-225, 1993.
- [7] Mack, T. *The Standard Error of Chain Ladder Reserve Estimates: Recursive Calculation and Inclusion of a Tail Factor*, ASTIN Bulletin, **29**, 361-366, 1999.
- [8] Mack, T., Quarg, G., Braun C., *The Mean Square Error of Prediction in the Chain Ladder Reserving Method – A Comment*, ASTIN Bulletin, **36**, 543-552, 2006.
- [9] G.C. Taylor, *Claims Reserving in Non-Life Insurance*, North-Holland 1986.
- [10] Wüthrich, M., Merz, M., *Stochastic claims reserving methods in insurance*, Wiley, 2008.
- [11] Wüthrich, M., Merz, M., *Modelling the claims development result for solvency purposes*, CAS E-Forum, Fall 2008, 542-568.



# QUADERNI DEL DIPARTIMENTO DI ECONOMIA, FINANZA E STATISTICA

Università degli Studi di Perugia

1	Gennaio 2005	Giuseppe CALZONI Valentina BACCHETTINI	Il concetto di competitività tra approccio classico e teorie evolutive. Caratteristiche e aspetti della sua determinazione
2	Marzo 2005	Fabrizio LUCIANI Marilena MIRONIUC	Ambiental policies in Romania. Tendencies and perspectives
3	Aprile 2005	Mirella DAMIANI	Costi di agenzia e diritti di proprietà: una premessa al problema del governo societario
4	Aprile 2005	Mirella DAMIANI	Proprietà, accesso e controllo: nuovi sviluppi nella teoria dell'impresa ed implicazioni di corporate governance
5	Aprile 2005	Marcello SIGNORELLI	Employment and policies in Europe: a regional perspective
6	Maggio 2005	Cristiano PERUGINI Paolo POLINORI Marcello SIGNORELLI	An empirical analysis of employment and growth dynamics in the italian and polish regions
7	Maggio 2005	Cristiano PERUGINI Marcello SIGNORELLI	Employment differences, convergences and similarities in italian provinces
8	Maggio 2005	Marcello SIGNORELLI	Growth and employment: comparative performance, convergences and co-movements
9	Maggio 2005	Flavio ANGELINI Stefano HERZEL	Implied volatilities of caps: a gaussian approach
10	Giugno 2005	Slawomir BUKOWSKI	EMU – Fiscal challenges: conclusions for the new EU members
11	Giugno 2005	Luca PIERONI Matteo RICCIARELLI	Modelling dynamic storage function in commodity markets: theory and evidence
12	Giugno 2005	Luca PIERONI Fabrizio POMPEI	Innovations and labour market institutions: an empirical analysis of the Italian case in the middle 90's
13	Giugno 2005	David ARISTEI Luca PIERONI	Estimating the role of government expenditure in long-run consumption
14	Giugno 2005	Luca PIERONI Fabrizio POMPEI	Investimenti diretti esteri e innovazione in Umbria
15	Giugno 2005	Carlo Andrea BOLLINO Paolo POLINORI	Il valore aggiunto su scala comunale: la Regione Umbria 2001-2003
16	Giugno 2005	Carlo Andrea BOLLINO Paolo POLINORI	Gli incentivi agli investimenti: un'analisi dell'efficienza industriale su scala geografica regionale e sub regionale

17	Giugno 2005	Antonella FINIZIA Riccardo MAGNANI Federico PERALI Paolo POLINORI Cristina SALVIONI	Construction and simulation of the general economic equilibrium model Meg-Ismea for the Italian economy
18	Agosto 2005	Elżbieta KOMOSA	Problems of financing small and medium-sized enterprises. Selected methods of financing innovative ventures
19	Settembre 2005	Barbara MROCZKOWSKA	Regional policy of supporting small and medium-sized businesses
20	Ottobre 2005	Luca SCRULLA	Clustering multivariate spatial data based on local measures of spatial autocorrelation
21	Febbraio 2006	Marco BOCCACCIO	Crisi del welfare e nuove proposte: il caso dell'unconditional basic income
22	Settembre 2006	Mirko ABBRITTI Andrea BOITANI Mirella DAMIANI	Unemployment, inflation and monetary policy in a dynamic New Keynesian model with hiring costs
23	Settembre 2006	Luca SCRULLA	Subset selection in dimension reduction methods
24	Ottobre 2006	Sławomir I. BUKOWSKI	The Maastricht convergence criteria and economic growth in the EMU
25	Ottobre 2006	Jan L. BEDNARCZYK	The concept of neutral inflation and its application to the EU economic growth analyses
26	Dicembre 2006	Fabrizio LUCIANI	Sinossi dell'approccio teorico alle problematiche ambientali in campo agricolo e naturalistico; il progetto di ricerca nazionale F.I.S.R. – M.I.C.E.N.A.
27	Dicembre 2006	Elvira LUSSANA	Mediterraneo: una storia incompleta
28	Marzo 2007	Luca PIERONI Fabrizio POMPEI	Evaluating innovation and labour market relationships: the case of Italy
29	Marzo 2007	David ARISTEI Luca PIERONI	A double-hurdle approach to modelling tobacco consumption in Italy
30	Aprile 2007	David ARISTEI Federico PERALI Luca PIERONI	Cohort, age and time effects in alcohol consumption by Italian households: a double-hurdle approach
31	Luglio 2007	Roberto BASILE	Productivity polarization across regions in Europe
32	Luglio 2007	Roberto BASILE Davide CASTELLANI Antonello ZANFEI	Location choices of multinational firms in Europe: the role of EU cohesion policy
33	Agosto 2007	Flavio ANGELINI Stefano HERZEL	Measuring the error of dynamic hedging: a Laplace transform approach

34	Agosto 2007	Stefano HERZEL Cătălin STĂRICĂ Thomas NORD	The IGARCH effect: consequences on volatility forecasting and option trading
35	Agosto 2007	Flavio ANGELINI Stefano HERZEL	Explicit formulas for the minimal variance hedging strategy in a martingale case
36	Agosto 2007	Giovanni BIGAZZI	The role of agriculture in the development of the people's Republic of China
37	Settembre 2007	Enrico MARELLI Marcello SIGNORELLI	Institutional change, regional features and aggregate performance in eight EU's transition countries
38	Ottobre 2007	Paolo NATICCHIONI Andrea RICCI Emiliano RUSTICHELLI	Wage structure, inequality and skill-biased change: is Italy an outlier?
39	Novembre 2007	The International Study Group on Exports and Productivity	Exports and productivity. Comparable evidence for 14 countries
40	Dicembre 2007	Gaetano MARTINO Paolo POLINORI	Contracting food safety strategies in hybrid governance structures
41	Dicembre 2007	Floro Ernesto CAROLEO Francesco PASTORE	The youth experience gap: explaining differences across EU countries
42	Gennaio 2008	Melisso BOSCHI Luca PIERONI	Aluminium market and the macroeconomy
43	Febbraio 2008	Flavio ANGELINI Marco NICOLSI	Hedging error in Lévy models with a fast Fourier Transform approach
44	Febbraio 2008	Luca PIERONI Giorgio d'AGOSTINO Marco LORUSSO	Can we declare military Keynesianism dead?
45	Febbraio 2008	Pierluigi GRASSELLI Cristina MONTESI Paola IANNONE	Mediterranean models of Welfare towards families and women
46	Marzo 2008	Mirella DAMIANI Fabrizio POMPEI	Mergers, acquisitions and technological regimes: the European experience over the period 2002-2005
47	Marzo 2008	Bruno BRACALENTE Cristiano PERUGINI	The Components of Regional Disparities in Europe
48	Marzo 2008	Cristiano PERUGINI Fabrizio POMPEI Marcello SIGNORELLI	FDI, R&D and Human Capital in Central and Eastern European Countries
49	Marzo 2008	Cristiano PERUGINI	Employment and Unemployment in the Italian Provinces
50	Marzo 2008	Sławomir I. BUKOWSKI	On the road to the euro zone. Currency rate stabilization: experiences of the selected EU countries
51	Aprile 2008	Bruno BRACALENTE Cristiano PERUGINI Fabrizio POMPEI	Homogeneous, Urban Heterogeneous, or both? External Economies and Regional Manufacturing Productivity in Europe

52	Aprile 2008	Gaetano MARTINO Cristiano PERUGINI	Income inequality within European regions: determinants and effects on growth
53	Aprile 2008	Jan L. BEDNARCZYK	Controversy over the interest rate theory and policy. Classical approach to interest rate and its continuations
54	Aprile 2008	Bruno BRACALENTE Cristiano PERUGINI	Factor decomposition of cross-country income inequality with interaction effects
55	Aprile 2008	Cristiano PERUGINI	Employment Intensity of Growth in Italy. A Note Using Regional Data
56	Aprile 2008	Cristiano PERUGINI Fabrizio POMPEI	Technological Change, Labour Demand and Income Distribution in European Union Countries
57	Aprile 2008	Simona BIGERNA Paolo POLINORI	L'analisi delle determinanti della domanda di trasporto pubblico nella città di Perugia
58	Maggio 2008	Simona BIGERNA Paolo POLINORI	The willingness to pay for Renewable Energy Sources (RES): the case of Italy with different survey approaches and under different EU "climate vision". First results
59	Giugno 2008	Simona BIGERNA Paolo POLINORI	Ambiente operativo ed efficienza nel settore del Trasporto Pubblico Locale in Italia
60	Ottobre 2008	Pierluigi GRASSELLI Cristina MONTESI Roberto VIRDI	L'interpretazione dello spirito del dono
61	Novembre 2008	Antonio BOGGIA Fabrizio LUCIANI Gianluca MASSEI Luisa PAOLOTTI	L'impatto ambientale ed economico del cambiamento climatico sull'agricoltura
62	Novembre 2008	Elena STANGHELLINI Francesco Claudio STINGO Rosa CAPOBIANCO	On the estimation of a binary response model in a selected population
63	Dicembre 2008	Gianna FIGÀ-TALAMANCA	Limit results for discretely observed stochastic volatility models with leverage effect
64	Maggio 2009	Mirella DAMIANI Andrea RICCI	Factors behind performance-related pay: evidence from Italy
65	Giugno 2009	Alessandra RIGHI Dario SCIULLI	The Timing of the School-to-Permanent Work Transition: a Comparison across Ten European Countries
66	Settembre 2009	Fabrizio LUCIANI	Economia agraria e pianificazione economica territoriale nel Parco nazionale del Sagarmatha (Everest, Nepal)
67	Settembre 2009	Valentina TIECCO	I regimi di protezione dell'impiego

68	Ottobre 2009	Gianna FIGÀ-TALAMANCA	Path properties of simulation schemes for the Heston stochastic volatility model
69	Ottobre 2009	Cristina MONTESI	A comparative analysis of different business ethics in the perspective of the Common Good
70	Ottobre 2009	Luisa FRANZINI Margherita GIANNONI	Determinants of Health Disparities in Italian Regions
71	Novembre 2009	Flavio ANGELINI Stefano HERZEL	Evaluating Discrete Dynamic Strategies in Affine Models
72	Novembre 2009	Giuseppe ARBIA Michele BATTISTI Gianfranco DI VAIO	Institutions and geography: Empirical test of spatial growth models for European regions
73	Gennaio 2010	Mirella DAMIANI Andrea RICCI	Performance-Related Pay, Unions and Productivity in Italy: evidence from quantile regressions
74	Febbraio 2010	Davide CASTELLANI Fabio PIERI	The Effect of Foreign Investments on European Regional Productivity
75	Luglio 2010	Guglielmo M. CAPORALE Davide CIFERRI Alessandro GIRARDI	Time-varying spot and futures oil price dynamics
76	Settembre 2010	Mirella DAMIANI	Labour regulation, corporate governance and varieties of capitalism
77	Settembre 2010	Dario SCIULLI Marcello SIGNORELLI	University-to-work transitions: the case of Perugia
78	Ottobre 2010	Olga DEMIDOVA Marcello SIGNORELLI	The Impact of Crises on Youth Unemployment of Russian Regions: An Empirical Analysis
79	Ottobre 2010	Misbah T. CHOUDHRY Enrico MARELLI Marcello SIGNORELLI	The Impact of Financial Crises on Youth Unemployment Rate
80	Novembre 2010	Marco BELLUCCI	Fusioni ed acquisizioni: determinanti ed effetti in un confronto europeo
81	Dicembre 2010	Silvia MICHELI	Learning Curve and Wind Power
82	Dicembre 2010	Leonardo BECCHETTI Stefano CASTRIOTA Elena GIACHIN RICCA	Beyond the Joneses: inter-country income comparisons and happiness
83	Gennaio 2011	Davide CASTELLANI Fabio PIERI	Foreign Investments and Productivity Evidence from European Regions
84	Febbraio 2011	Stefano HERZEL Marco NICOLOSI Cătălin STĂRICĂ	The cost of sustainability on optimal portfolio choices
85	Marzo 2011	Pierluigi GRASSELLI Cristina MONTESI	Politiche orientate al bene comune e politiche attive del lavoro
86	Marzo 2011	Massimo DE FELICE Franco MORICONI	Un'estensione stocastica del modello "Fisher-Lange"

**I QUADERNI DEL DIPARTIMENTO DI ECONOMIA**  
**Università degli Studi di Perugia**

<b>1</b>	Dicembre 2002	Luca PIERONI:	Further evidence of dynamic demand systems in three european countries
<b>2</b>	Dicembre 2002	Luca PIERONI Paolo POLINORI:	Il valore economico del paesaggio: un'indagine microeconomica
<b>3</b>	Dicembre 2002	Luca PIERONI Paolo POLINORI:	A note on internal rate of return
<b>4</b>	Marzo 2004	Sara BIAGINI:	A new class of strategies and application to utility maximization for unbounded processes
<b>5</b>	Aprile 2004	Cristiano PERUGINI:	La dipendenza dell'agricoltura italiana dal sostegno pubblico: un'analisi a livello regionale
<b>6</b>	Maggio 2004	Mirella DAMIANI:	Nuova macroeconomia keynesiana e quasi razionalità
<b>7</b>	Maggio 2004	Mauro VISAGGIO:	Dimensione e persistenza degli aggiustamenti fiscali in presenza di debito pubblico elevato
<b>8</b>	Maggio 2004	Mauro VISAGGIO:	Does the growth stability pact provide an adequate and consistent fiscal rule?
<b>9</b>	Giugno 2004	Elisabetta CROCI ANGELINI Francesco FARINA:	Redistribution and labour market institutions in OECD countries
<b>10</b>	Giugno 2004	Marco BOCCACCIO:	Tra regolamentazione settoriale e antitrust: il caso delle telecomunicazioni
<b>11</b>	Giugno 2004	Cristiano PERUGINI Marcello SIGNORELLI:	Labour market performance in central european countries
<b>12</b>	Luglio 2004	Cristiano PERUGINI Marcello SIGNORELLI:	Labour market structure in the italian provinces: a cluster analysis
<b>13</b>	Luglio 2004	Cristiano PERUGINI Marcello SIGNORELLI:	I flussi in entrata nei mercati del lavoro umbri: un'analisi di cluster
<b>14</b>	Ottobre 2004	Cristiano PERUGINI:	Una valutazione a livello microeconomico del sostegno pubblico di breve periodo all'agricoltura. Il caso dell'Umbria attraverso i dati RICA-INEA
<b>15</b>	Novembre 2004	Gaetano MARTINO Cristiano PERUGINI	Economic inequality and rural systems: empirical evidence and interpretative attempts
<b>16</b>	Dicembre 2004	Federico PERALI Paolo POLINORI Cristina SALVIONI Nicola TOMMASI Marcella VERONESI	Bilancio ambientale delle imprese agricole italiane: stima dell'inquinamento effettivo